Hydrological parameter estimations from a conservative tracer test with variable-density effects at the Boise Hydrogeophysical Research Site

B. Dafflon,[1,2] W. Barrash,[1] M. Cardiff,[1] and T. C. Johnson[3]

Received 13 April 2011; revised 7 October 2011; accepted 27 October 2011; published 15 December 2011.

[1] Reliable predictions of groundwater flow and solute transport require an estimation of the detailed distribution of the parameters (e.g., hydraulic conductivity, effective porosity) controlling these processes. However, such parameters are difficult to estimate because of the inaccessibility and complexity of the subsurface. In this regard, developments in parameter estimation techniques and investigations of field experiments are still challenging and necessary to improve our understanding and the prediction of hydrological processes. Here we analyze a conservative tracer test conducted at the Boise Hydrogeophysical Research Site in 2001 in a heterogeneous unconfined fluvial aquifer. Some relevant characteristics of this test include: variable-density (sinking) effects because of the injection concentration of the bromide tracer, the relatively small size of the experiment, and the availability of various sources of geophysical and hydrological information. The information contained in this experiment is evaluated through several parameter estimation approaches, including a grid-search-based strategy, stochastic simulation of hydrological property distributions, and deterministic inversion using regularization and pilot-point techniques. Doing this allows us to investigate hydraulic conductivity and effective porosity distributions and to compare the effects of assumptions from several methods and parameterizations. Our results provide new insights into the understanding of variable-density transport processes and the hydrological relevance of incorporating various sources of information in parameter estimation approaches. Among others, the variable-density effect and the effective porosity distribution, as well as their coupling with the hydraulic conductivity structure, are seen to be significant in the transport process. The results also show that assumed prior information can strongly influence the estimated distributions of hydrological properties.

Citation: Dafflon, B., W. Barrash, M. Cardiff, and T. C. Johnson (2011), Hydrological parameter estimations from a conservative tracer test with variable-density effects at the Boise Hydrogeophysical Research Site, *Water Resour. Res.*, *47*, W12513, doi:10.1029/2011WR010789.

1. Introduction

[2] Accurate prediction of flow and solute transport in aquifers is a key prerequisite for effective sustainable management and remediation of groundwater resources. In fact, such predictions are very challenging because of the complexity of the hydrological processes and the difficulty in obtaining information about important flow parameters in the subsurface. The general procedure for investigating the distribution of hydrological properties, such as hydraulic conductivity (K) and effective porosity (ϕ), for example, includes (1) experimental design, (2) data acquisition, (3) model conceptualization, and (4) parameter estimation [e.g., *de Marsily et al.*, 1999; *Carrera et al.*, 2005]. First, designing an experiment involves the development of data acquisition techniques and strategies that are sensitive to properties and processes of interest. Such strategies also need to be optimized with regard to actual field and boundary conditions as well as cost and time factors. Second, the objective of data acquisition is to gain as much reliable and useful information as possible for the goals of a given study. Third, model conceptualization includes identifying the nature of the problem and hydrogeologic system, selecting the governing equations, defining the boundary conditions and the time regime, selecting the information sources of interest, and the way of integrating them. This step is highly dependent on the prior understanding of the system and processes. Fourth and finally, the parameter estimation is focused on assigning values to the model properties identified as controlling the processes of interest. This step is often performed through inverse modeling or simulation approaches to infer distributions of properties that allow matching (to an acceptable tolerance) of measurements of the processes of interest.

[1] Center for Geophysical Investigation of the Shallow Subsurface, Boise State University, Boise, Idaho, USA.
[2] Now at Lawrence Berkeley National Laboratory, Berkeley, California, USA.
[3] Pacific Northwest National Laboratory, Richland, Washington, USA.

0043-1397/11/2011WR010789

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2011		**00-00-2011 to 00-00-2011**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Hydrological Parameter Estimations From A Conservative Tracer Test With Variable-Density Effects At The Boise Hydrogeophysical Research Site	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Boise State University,Center for Geophysical Investigation of the Shallow Subsurface,Boise,ID,83725**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Water Resources Research, Vol. 47, 2011,Government or Federal Purpose Rights License.

14. ABSTRACT
Reliable predictions of groundwater flow and solute transport require an estimation of the detailed distribution of the parameters (e.g., hydraulic conductivity, effective porosity)controlling these processes. However, such parameters are difficult to estimate because of the inaccessibility and complexity of the subsurface. In this regard, developments in parameter estimation techniques and investigations of field experiments are still challenging and necessary to improve our understanding and the prediction of hydrological processes. Here we analyze a conservative tracer test conducted at the Boise Hydrogeophysical Research Site in 2001 in a heterogeneous unconfined fluvial aquifer. Some relevant characteristics of this test include: variable-density (sinking) effects because of the injection concentration of the bromide tracer, the relatively small size of the experiment, and the availability of various sources of geophysical and hydrological information. The information contained in this experiment is evaluated through several parameter estimation approaches, including a grid-search-based strategy, stochastic simulation of hydrological property distributions, and deterministic inversion using regularization and pilot-point techniques.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**20**	

[3] Advancements in the above areas are supported by numerical, laboratory, and field experiments [e.g., *Rubin*, 1995; *Yeh and Liu*, 2000; *Cardiff et al.*, 2009; *Pollock and Cirpka*, 2010; *Sudicky et al.*, 2010]. In particular, field experiments in various hydrological and geological contexts, at various scales, and with various survey strategies, are essential to better understand flow and transport processes, validate and improve modeling approaches, and guide novel developments [e.g., *Sudicky*, 1986; *Ptak and Schmid*, 1996; *Illman et al.*, 2009; *Cvetkovic et al.*, 2010; *Brauchler et al.*, 2011]. Such field experiments commonly include variations on tracer tests, multiwell pumping or injection tests, and single-well slug or flowmeter tests or direct-push profiles [e.g., *Hyndman and Gorelick*, 1996; *Brauchler et al.*, 2003; *Doherty*, 2003; *Zhu and Yeh*, 2005; *Cardiff et al.*, 2011]. In this study we present the analysis of a tracer test performed in 2001 at the Boise Hydrogeophysical Research Site (BHRS) [*Barrash et al.*, 2002; *Hausrath et al.*, 2002], and investigate a number of issues related to the transport of solute subject to variable-density effects in a heterogeneous unconfined aquifer at a local scale with several parameter estimation approaches.

[4] One main objective of studies about solute transport [e.g., *Mackay et al.*, 1986; *Sudicky*, 1986; *Leblanc et al.*, 1991; *Hyndman et al.*, 2000; *Vereecken et al.*, 2000; *Hubbard et al.*, 2001; *Scheibe and Chien*, 2003; *Müller et al.*, 2010; *Sudicky et al.*, 2010] is to improve the estimation of spatially variable properties controlling flow and transport, primarily hydraulic conductivity (K) and effective porosity (ϕ), and to improve the understanding and prediction of processes, such as transport, dispersion, and chemical and biological reactions among others [e.g., *de Marsily et al.*, 2005]. In this study, we evaluate the spatial distribution of hydrological properties (e.g., K, ϕ, dispersivity) controlling the observed transport process, the importance of each one in predicting this process, and the possible coupling effects between properties. The resulting understanding is useful not only to improve predictions or solve issues at one specific site, but also to allow the extrapolation of gained knowledge to other areas where only limited information may be available [e.g., *Hyndman et al.*, 2000; *Linde et al.*, 2006; *Dafflon et al.*, 2010].

[5] Another subject of interest in this study is how we can use indirect or supplemental information to constrain the distribution of hydrological properties controlling flow and transport processes and thus improve flow and transport predictions. In this context, geophysical methods have shown much potential [e.g., *Hyndman and Gorelick*, 1996; *Slater et al.*, 1997; *Hubbard et al.*, 2001; *Singha and Gorelick*, 2006; *Müller et al.*, 2010] because they provide a scale of spatial resolution and degree of subsurface coverage not available with traditional hydrological measurement techniques (e.g., borehole logs, core analyses, pumping, and tracer tests). In addition, geophysical methods are able to provide large quantities of informative data in less time and at lower cost than most hydrological methods. The BHRS is particularly suitable for investigating the value of supplemental information due to the broad database of geophysical and hydrological information available at this site. In this study we examine how hydrological information gained from a tracer test alone compares with information inferred from geophysical data, and how the ϕ distribution estimated from geophysical data can influence the hydrological parameter estimation process. This is essential to evaluate how geophysical data may be sensitive to hydrological properties, how such properties may be related together, and what can be expected from various hydrological, geophysical, and hydrogeophysical data sets.

[6] In addition, the use of parameter estimation techniques for variable-density solute transport studies are still computationally intensive and have not been investigated much with regard to the information content of hydrological data sets influenced by density-driven processes. Although a number of studies have recognized or described sinking of a tracer because of density effects [e.g., *Leblanc et al.*, 1991; *Vereecken et al.*, 2000; *Beinhorn et al.*, 2005; *Müller et al.*, 2010], the use of parameter estimation methods for problems that include density effects is still rare. Indeed, it has been noted that a significant number of studies neglect the variable-density effect although they should not [*Simmons*, 2005]. This is probably because of the increased complexity in modeling such processes and in processing and inverting such data, to significantly increased demand for computing power, and to increased difficulty in interpretation of the results. Here for inverting such data sets, we investigate several parameter estimation approaches, and some crucial issues such as the optimization of computing time and model size, the objective function used, and the implementation of inversion parametric constraints (i.e., "prior" information).

[7] This paper is organized as follows. First, the BHRS and hydrogeological and geophysical information relevant for the tracer test are presented. Then the tracer test is described as well as the parameters used for the flow and solute transport modeling. Next, the measured concentrations during the tracer test are used to evaluate the distribution of hydrological properties such as K and ϕ, and also to investigate the information contained in such measurements through various strategies including (1) a grid-search-based approach to infer relatively simple models of hydrological properties, (2) simulations of hydrological property distributions to consider the impact of introducing small-scale heterogeneity for predicting the measured concentrations, and (3) a deterministic inversion involving regularization and pilot-point techniques to evaluate various parameterizations of the problem.

2. Field Site Hydrogeology

[8] The BHRS is a hydrological and geophysical field research site located on a gravel bar adjacent to the Boise River ~15 km from downtown Boise, Idaho. The shallow, unconfined aquifer at the site consists of coarse unconsolidated fluvial deposits that are ~20-m-thick, have minimal fractions of silt and clay, and are underlain by a layer of red clay [*Barrash and Clemo*, 2002; *Barrash and Reboulet*, 2004]. At this site, 18 wells were emplaced carefully using a core drive-drill method to minimize the disturbance of the surrounding formation and were fully screened through the aquifer with 4-in PVC. The well field consists of 13 wells in the central area (~20 m in diameter) and five boundary wells ~10–35 m from the central area (Figure 1). The arrangement of the 13 inner wells is a central well (A1) surrounded by two concentric rings of six wells each

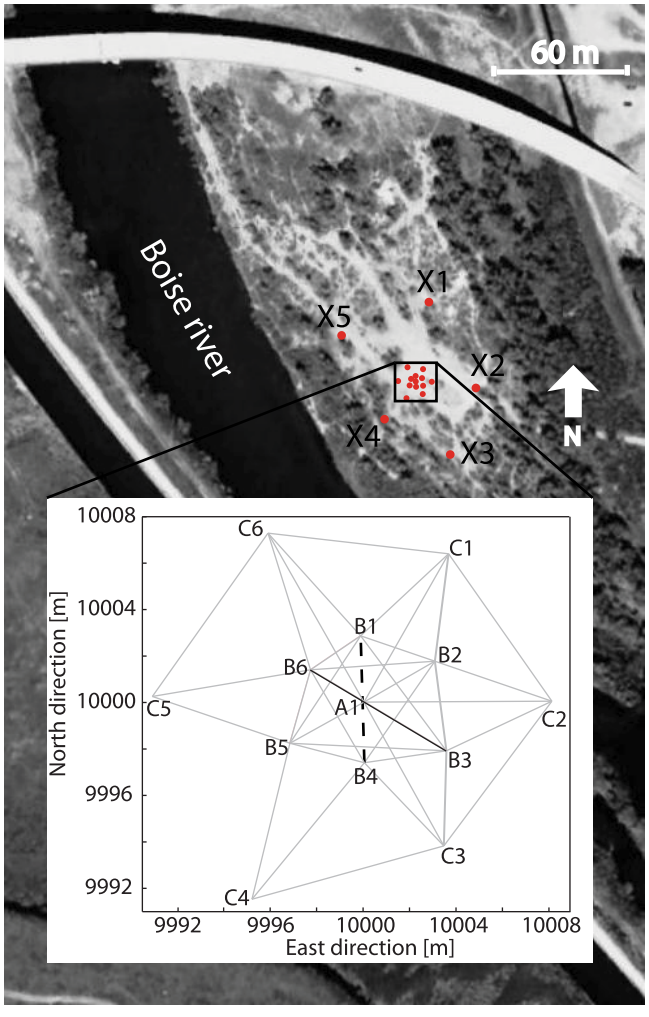

Figure 1. Aerial view of the BHRS and detailed map of the central area wells with lines to indicate between which wells the tracer test (black, solid) and static cross-hole GPR (gray) acquisitions have been performed. The time-lapse cross-hole GPR profile shown in this study is indicated with a dotted line.

(B1–B6 and C1–C6). The distances between the adjacent wells vary between 2.6 and 9.7 m. Their total depths vary between 18.2 and 20 m below the land surface which occurs between 849.32 m and 849.64 m above mean sea level (AMSL) in the central area.

[9] Some key structural information about the hydrogeological units and properties at this site has been obtained from neutron (porosity) logs and cross-hole ground-penetrating radar (GPR) data. The neutron-log data were measured every 0.06 m in each borehole and ϕ values were obtained from the measured count rate through a petrophysical transform [*Barrash and Clemo*, 2002]. At the BHRS, neutron-log data are primarily sensitive to total porosity in the water-saturated, unconsolidated, coarse (high-energy) sedimentary deposits that have been documented to have virtually no silt and clay [*Barrash and Reboulet*, 2004]. That is, the measured porosity values can be taken as equivalent to effective porosity values for this aquifer with the risk of only very limited overestimation. On the basis of these ϕ logs, *Barrash and Clemo* [2002] identified five ϕ units with distinct spatial distribution, mean, and variance (Figure 2). The fitted Gaussian functions to the ϕ distributions of units 1–5 have means equal to 0.18, 0.24, 0.172, 0.224, and 0.425, respectively, and standard deviations equal to 0.022, 0.038, 0.024, 0.05, and 0.055, respectively. Unit 5 is a high ϕ channel that thickens toward the Boise River and pinches out in the center of the well field. Units 1–4 are a sequence of conglomerates with gravel and cobble framework and sand to pebble matrix in the interstices of the framework. More recently, electrical capacitive conductivity measurements have allowed the identification of a subunit 2b, which is present in all of the measured A–C wells shown in Figure 1 except wells B1, B3, C1, and C2 [*Mwenifumbo et al.*, 2009]. The contacts between units 1, 2a-2b, 3, 4, and 5 are projected between wells B6, A1, and B3 in Figure 2.

[10] Numerous cross-hole GPR measurements have also been acquired at the BHRS, and from these, 31 intersecting cross-hole GPR profiles have been inverted together to obtain a highly resolved and internally consistent model of radar velocity distribution along the various profile directions [*Dafflon et al.*, 2011]. The obtained velocity distribution generally correlates well with complementary ϕ log data, with correlation coefficients varying between -0.32 and -0.79 over the site and with a mean equal to -0.57. Moreover, such GPR velocity tomograms are an important source of ϕ structural information between the boreholes

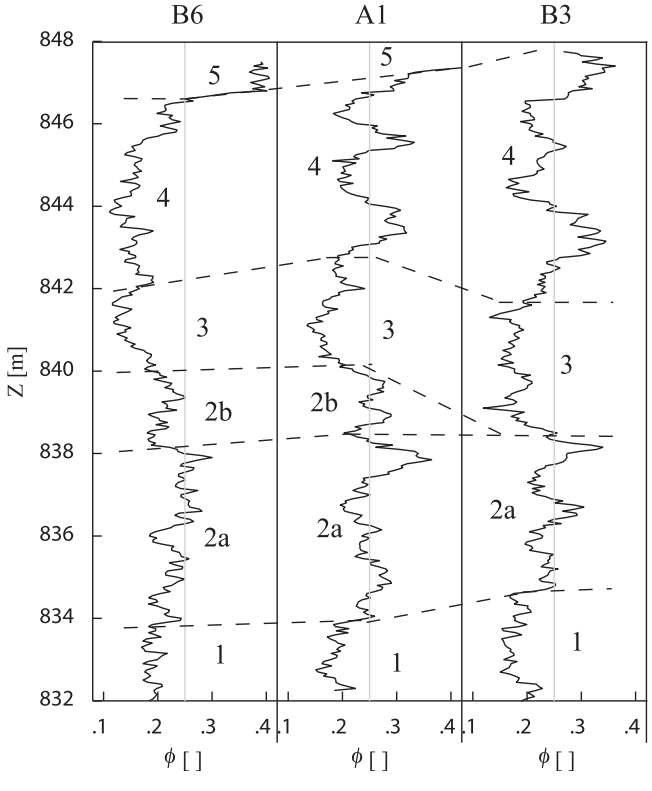

Figure 2. ϕ logs in wells B6, A1, and B3. For consideration of the general hydrogeological structure at the BHRS, contacts between units 1–5 are indicated based on information contained in the logs (shown here). Relative horizontal positions of well logs are not to scale.

because of their continuous spatial coverage and the strong petrophysical relationship (negative correlation) between GPR velocity and ϕ in the aquifer at the BHRS (except for unit 2b). Figure 3a shows the GPR velocity tomogram along the profile B6–B3 from the inversion of 31 profiles together. The imaged structures correlate well with the ϕ log data (Figure 2). Also, in Figure 3a, units 1 and 3 show relatively low internal variability, whereas units 2a and 4 show higher variability and the presence of smaller scale structures (lenses). This is in agreement with the previous geostatistical work of *Barrash and Clemo* [2002], which found that the variance in ϕ in units 2 and 4 (see values above) is ~1.5–2 times larger than that in units 1 and 3. We also note that the GPR clearly imaged the boundary between units 2a and 2b and thus corroborates the result obtained from electrical capacitive conductivity measurements that indicate unit 2 can be subdivided into units with different ϕ-electric/dielectric petrophysics [*Mwenifumbo et al.*, 2009].

[11] Finally, 3-D realizations of the ϕ distribution at the BHRS have been generated from the information contained in the multidirectional GPR velocity model and the ϕ logs (B. Dafflon and W. Barrash, 3-D stochastic estimation of porosity distribution: Benefits of using GPR velocity tomograms in simulated-annealing-based or Bayesian sequential simulation approaches, submitted to *Water Resources Research*, 2011) through a simulated annealing approach [*Dafflon et al.*, 2009]. In short, this method allows the successful fusion of the relatively large-scale information contained in the GPR velocity model and the smaller-scale information contained in the ϕ logs. To do this, a Monte Carlo-based process is used to optimize the structural variability of a realization (i.e., given a geostatistical target function) by sequentially and randomly perturbing this realization, where each cell is conditioned on the information available at that cell. This means that, in addition to the available GPR velocity and ϕ data, this approach uses the conditional relation inferred between these properties along the wells and with geostatistical functions obtained from the data. This approach has been shown to provide realistic 3-D simulations and to increase the accuracy of ϕ prediction compared to not using the GPR data (Dafflon and Barrash, submitted manuscript, 2011). Thus, the best 3-D ϕ estimates now available at the BHRS are obtained through the coupling of the ϕ logs and cross-hole GPR velocity data. One representative realization of the ϕ structure is shown in Figure 3b along the profile B6–B3. It is evident by inspection that the spatial variability and the ϕ values in each unit are very consistent with the ϕ data (Figure 2) and the GPR velocity imaged structures (Figure 3a).

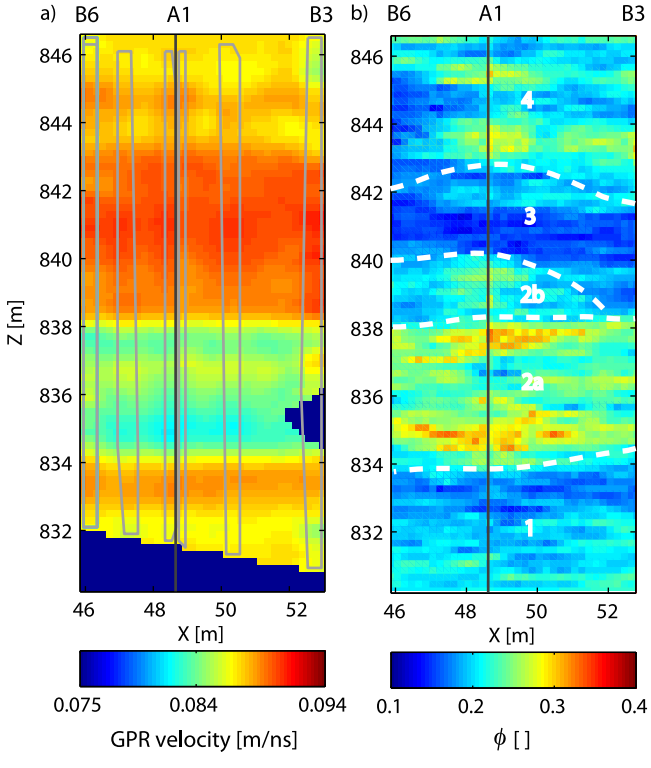

Figure 3. (a) Velocity tomogram for profile B3–B6 that was obtained by inverting together 31 cross-hole GPR data sets [*Dafflon et al.*, 2011]. Intersections with other profiles are delineated (gray). (b) ϕ distribution along profile B3–B6, extracted from a 3-D realization obtained from the ϕ logs and velocity tomograms using a simulated-annealing conditional simulation approach, shown with the main lithological unit boundaries identified at the BHRS. This realization can be considered as representative of a large number of realizations with regard to implications for this study.

3. Tracer Test

[12] The tracer test considered in this work was conducted at the BHRS in August 2001 [*Barrash et al.*, 2002]. A conservative tracer was injected into a 4-m interval that spanned the boundary between units 2a and 3 in well B3, and was removed by pumping at well B6 from a 4-m interval located at the same elevation as the injection interval (Figure 4). Wells B3 and B6 are 6.9 m apart and their alignment is oriented ~30° west of the natural head gradient direction. A third well (A1) is located between, and on the same alignment as wells B3 and B6 (Figure 1). Well A1 was equipped to collect samples from 20 0.25 m-long sampling zones isolated by packers and situated in a 5-m interval above 836.575 m elevation (Figure 4), and accessed by ports and dedicated tubing. The same protocol was used at wells B1, B2, B4, and B5 to collect samples from six 1-m-long zones centered on the 4-m interval of the injection zone, although these wells were on the margin of the expected plume path. Similarly, samples were taken from six 1-m-long intervals at B6 (pumping well), which included four 1-m-long intervals that had perforations in the riser to allow pumping intake over the injection zone's elevation span. Flowrate measurements were taken with a digital flowmeter for injection and pumping in wells B3 and B6, respectively. (See *Barrash et al.* [2002] and *Hausrath et al.* [2002] for additional details on well instrumentation, sampling protocols, measurements, and results for the tracer test.)

[13] The experiment started with the injection of the bromide tracer with a concentration of 7.56 g L^{-1} in well B3 for 33.3 min at a rate of 111.6 L min^{-1}. Approximately 35 min after the injection was completed, pumping was started from well B6 with a low rate of ~20 L min^{-1}. These tracer test operational parameters were designed based on prior numerical modeling [*Barrash et al.*, 2002]. The use of such

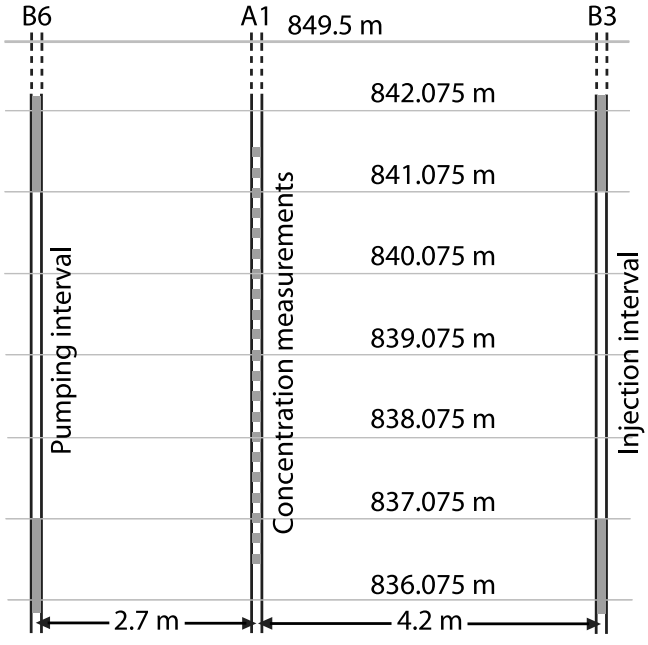

Figure 4. Geometry of the tracer test experiment performed in 2001 at the BHRS. Wells extend above to land surface at ~849.5 m elevation, and below to ~831.2 m elevation.

a pumping rate was intended to (1) assure the plume passage through A1, (2) ensure that the plume would remain within the ring of B wells, (3) completely remove the tracer from the field, and (4) minimize the influence of pumping on the interval between wells B3 and A1. Although it was planned to pump continuously at the same rate for the whole duration of the test, some variations occurred due to two power failures, cessation of pumping for GPR tomography, and some minor drift and fluctuation of the pumping rate [see *Barrash et al.*, 2002, Figure 34]. The main features of the pumping rate measured through the first 14.1 d are as follows: a rate equal to ~18.69 L min^{-1} between 35 min and 6.09 ds, then 17.66 L min^{-1} until day 14.1 with three interruptions: between 8.93 and 9.33 d, 11.81 and 12.09 d, and 13.81 and 14.1 d. Finally, after 14.1 d the pumping rate was increased to 98.42 L min^{-1} to ensure a nearly complete recovery of tracer from the aquifer. In this regard, >95% of bromide mass was recovered [*Hausrath et al.*, 2002].

[14] Groundwater samples of fluid for concentration measurements were collected with peristaltic pumps from the sampling zones in well A1 prior to the injection to get background concentration values, from the mixing tank prior to injection, from the injection line at 4 min intervals during injection, from the sampling zones in A1 and all B-wells except B3, and from the pumping line at B6 every 4 h after the injection, from 1 August until 16 August (see *Hausrath et al.* [2002] for additional detail on sampling and measurement). Analyses (temperature corrected) with an Accumet AR50 m and conductivity probe were completed at the test site within 4 hr of collection for all sample events during the test. The conductivity of the samples was measured and converted to bromide concentration with a linear relationship based on laboratory experiments [*Barrash et al.*, 2002; *Hausrath et al.*, 2002]. Subsequent independent laboratory tests at Boise State University and at a commercial laboratory demonstrated that conductivity measurements and bromide concentrations have a correlation coefficient of 0.994, and thus the conductivity measurements and the calibrated linear relationship provide close approximation to the bromide concentrations over the full range of concentrations observed during the tracer test [*Hausrath et al.*, 2002]. Breakthrough behavior at the 20 zones in A1 (Figure 5) indicates that these field and laboratory methods were able to detect small differences in tracer concentration between adjacent 0.25-m thick zones throughout the test. From Figure 5 it can be seen that the concentration breakthrough curves in well A1 can be classified into four vertically contiguous depth intervals based on similarities in the time of breakthrough and the overall magnitude and variation in the concentration measurements.

[15] Water level changes during the tracer experiment were mostly related to important changes in the injection and pumping rates at the site, which have been included here in the modeling. Furthermore, the flowrate and stage in the Boise River near the BHRS were relatively steady. Daily evapotranspiration cycles caused visible water level fluctuations in all of the surveyed wells with amplitudes ranging from 0.009 to 0.027 m. These influences on the flow and water level have not been considered in the hydrological model as they have been shown to not significantly influence the results [*Nelson*, 2007] and would unnecessarily

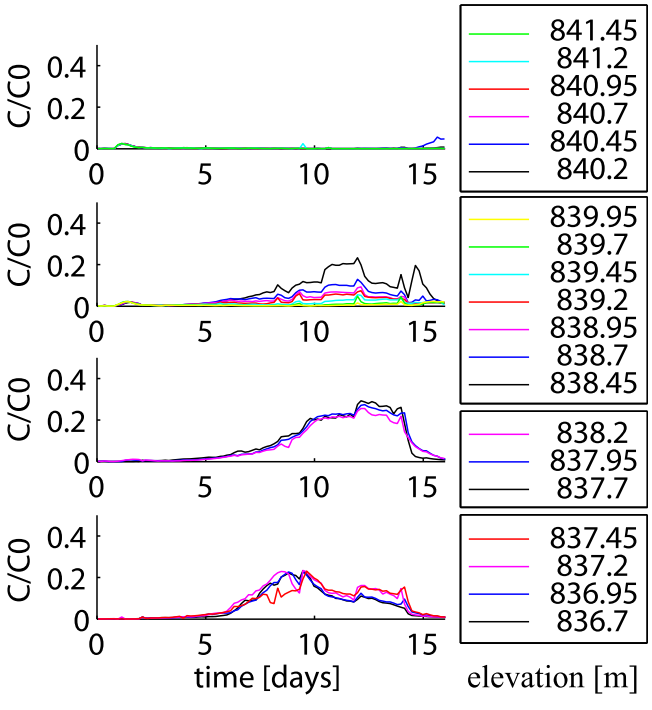

Figure 5. Tracer test breakthrough curves (BTCs) measured in 20 sampling intervals along well A1 (shown in Figure 4). Based on their general behaviors and shapes, the BTCs have been divided into four groups related to depth intervals where measurements were taken.

complicate the setting of the boundary conditions of the model. Head change measurements during the test were taken with vented strain-gage transducers (Keller model 730) in C and X wells, prototype fiber-optic transducers (Roctest MEMS-based Fabray-Perot type) in A and B wells, and e-tape measurements by hand in X wells (see *Barrash et al.* [2002] for additional details on head measurements during the tracer test). Drift after early time affected the prototype fiber optic transducers limiting quantitative measurements at the A and B wells. For this study we decided to use the manually measured water levels in X-wells only because they are the most accurate measurements ($\leq$0.003 m repeatability for absolute position), and they occur at our model boundary without imposing much influence on local hydrological or tracer plume behavior. In fact, these head measurements are fairly constant during each period of constant pumping rate (with a variation in the range of the daily evapotranspiration cycles described above) and show some more clear changes when the pumping rate changes significantly. In the modeling step, constant values have been set for each of these periods.

[16] Finally, over the course of the tracer test cross-hole GPR data were collected almost daily between B1–B4 and B2–B4, and on three separate days between B3–B6. For the purpose of this study we only considered the data along profile B1–B4 (transverse to the plume and intersecting well A1 with high-resolution tracer sampling control), that have been successfully processed to date [*Johnson et al.*, 2007] and constitute the most complete time-lapse survey. To obtain this profile, data were acquired with a 100 MHz Mala Geosciences RAMAC GPR system at 20 cm spacing in the receiver well B1 and 5 cm spacing in the transmitter well B4. *Johnson et al.* [2007] used these data with a Fresnel zone attenuations-difference tomography approach (FADT) to reconstruct bulk electrical conductivity changes in the subsurface over the course of the tracer test. Here Figure 6 shows the estimated bulk electrical conductivity difference distribution at various times along profile B1–B4 obtained by *Johnson et al.* [2007]. These time-lapse images are valuable for semiquantitative interpretation of the large-scale behavior of the tracer, the approximate depth reached by the sinking tracer, and the relative lateral spreading or positioning of the tracer.

4. Flow and Solute Transport Model Setup

[17] Based on results of initial modeling of the BHRS tracer test [*Nelson*, 2007], we decided to model this test using the following suite of model programs: MODFLOW [*Harbaugh et al.*, 2000], MT3DMS [*Zheng and Wang*, 1999; *Zheng*, 2005], and SEAWAT [e.g., *Guo and Langevin*, 2002; *Langevin et al.*, 2003, 2007], which are fully public-domain codes for 3-D flow and variable-density transport modeling. SEAWAT combines a modified version of MODFLOW and MT3DMS into a single program that solves the coupled flow and solute-transport equations to handle variable-density transport. This is done by iteratively solving the flow and transport governing equations and generating newly calculated heads based on the concept of an equivalent freshwater head in a saline groundwater environment. That is, at each step the density-dependent heads are transformed into freshwater heads using concentrations calculated in MT3DMS prior to solving the flow and transport equations again.

[18] It is important to note that accounting for the density effect is necessary for the 2001 BHRS tracer test [*Nelson*, 2007]. This is especially evident when looking at the shape and variations of the breakthrough curves (BTCs) at the different sampling intervals in A1 (Figure 5). Such sinking effects have also been observed in a number of similar field tracer experiments through well data and/or geophysical imaging [e.g., *Mackay et al.*, 1986; *Sudicky*, 1986; *Leblanc et al.*, 1991; *Vereecken et al.*, 2000; *Beinhorn et al.*, 2005; *Müller et al.*, 2010] and are discussed in review publications about variable-density flow issues [e.g., *Simmons et al.*, 2001; *Diersch and Kolditz*, 2002; *Beinhorn et al.*, 2005; *Simmons*, 2005]. In the present study, the bromide concentration was $\sim$7.5 g L^{-1} at the time of injection, or about one fifth of the concentration of salt in ocean water, which was shown to be very significant with regard to density effects [*Nelson*, 2007]. Enumeration of influences of tracer concentration on sinking and of concentrations used in various field experiments can be found in the work of *Beinhorn et al.* [2005]. Unfortunately, the inclusion of density effects in modeling such tracer experiments results in a significant increase in computing time of at least one order of magnitude. Although the transport can be simulated in a subdomain of the flow model (i.e., by identifying specified heads

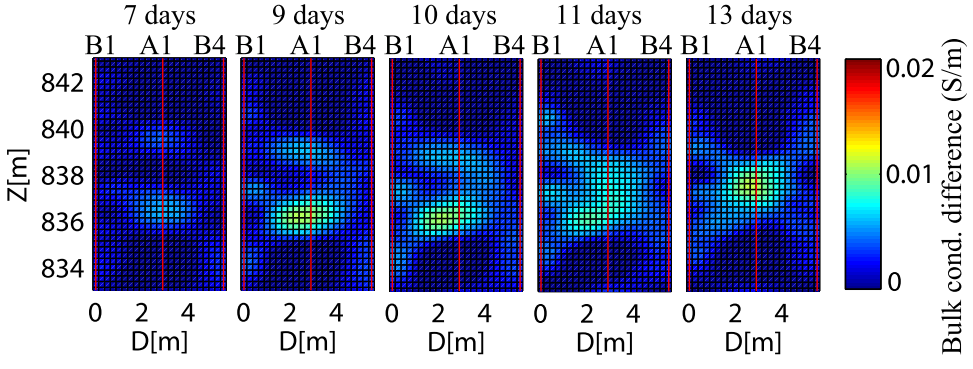

Figure 6. Bulk electrical conductivity difference along profile B1–B4 obtained from time-lapse GPR cross-hole measurements using a Fresnel zone attenuation-difference tomography approach [*Johnson et al.*, 2007].

along marginal flow lines) there is still a trade-off between increased model complexity (because of nonlinear density-dependent flow) and computing time.

[19] The model and boundary conditions were defined using prior information for this site and details of the tracer test. Boundaries for the unconfined aquifer model follow the X ring of wells (Figure 1) and thus are controlled by head measurements at the X-wells during the experiment. Given the relatively large distance between the B-wells and the X-wells, and the relatively small (observed) head variation in the X-wells (section 3), the specified heads at the X-wells do not significantly affect the flow and solute transport behavior locally in the region of tracer transport [*Nelson*, 2007]. The natural head gradient before the start of the experiment was ~0.001 in the flow direction oriented ~30° east of the B3–B6 alignment. Within the numerical model, cell dimensions are generally 0.2 m × 0.2 m × 0.2 m in the region of the B-wells (where the tracer transport occurs), and then progressively expand outward with a maximum size of 1 m × 1 m × 0.5 m. Relevant injection and pumping variations during the test are included in the model. The injection and pumping intervals were modeled with high K equal to 1 m s^{-1} and ϕ equal to 1 to simulate well hydraulics. Aquifer specific yield, specific storage, and the diffusion coefficient, which are relatively insensitive parameters in this experiment, are set to constants: Sy = 0.21, Ss = 4.5e-5 m^{-1}, and D = 1.34e-9 m^2 s^{-1}. K is treated as isotropic on the basis of prior information [*Barrash et al.*, 2006]. Also, one parameter that is included in SEAWAT and which influences density effects is the slope of the linear equation of state that relates fluid density to solute concentration [*Langevin et al.*, 2007]. This parameter, commonly set to 0.7143 for fresh- and saltwater interactions, is approximated by dividing the density difference over the range of end-member fluids by the difference in concentration between the end-member fluids [*Langevin et al.*, 2007]. This means that this parameter depends, in part, on chemical properties, pressure, and temperature of the injected tracer solution and the ambient water. Although this slope value is not discussed much in the literature and is relatively complex to estimate in a field experiment, preliminary tests have shown that changes in the result of this study are limited with regard to the uncertainty involved in this parameter. Numerical and laboratory study about this parameter would definitely be worthwhile.

[20] Finally, longitudinal, horizontal transverse, and vertical transverse dispersivities have been set to 0.06 m (+0.04, −0.02), 0.18 m (+0.07, −0.08), and 0.02 m (+0.03, −0.01), respectively, based on (1) grid search evaluations using a homogeneous model, (2) consideration of the shape of the simulated breakthrough curves (BTCs) compared to the measured BTCs, and (3) plume position information in GPR time-lapse images. The values in parentheses indicate ranges of dispersivity values that influence the simulated concentration only by slightly shifting absolute concentration values by about a proportional amount. Such variations have been shown to be similar to those that could be produced by error in positions within wells (i.e., deviation). One should note that, based on some additional tests and on the results of a study that included evaluating the effect of longitudinal and transverse dispersivity values on the shape of a plume [*Sudicky et al.*, 2010], the influence of varying dispersivity values to some degree (as above), does not significantly influence the shape of the plume. Although in this study the longitudinal dispersivity is smaller than the transverse dispersivity, which is not commonly observed, we kept these values because that dispersivity is not only simulating the spreading of the plume produced by nonmodeled heterogeneity but it is also a means of reproducing some effects of parameters that we did not include or failed to reproduce in the hydrological model [see also *Sudicky et al.*, 2010], such as (1) spatial trends of relatively lower heads further from the river associated with small river stage increases and/or evapotranspiration [*Malama and Johnson*, 2010; *Johnson*, 2011] that could produce transverse movement away from the river and thus perpendicular to the main tracer path direction, (2) possible northeastward-sloping topography on the unit 2–unit 3 contact in the northeastern portion of the B-ring of wells [*Barrash et al.*, 2002], and (3) a limitation in accuracy of modeling the hydrological processes and parameters of influence such as, for example, the coupling between spreading near injection zone and near-well hydrogeologic structures.

5. Parameter Estimation Strategies

[21] We estimated the distribution of K and ϕ from tracer BTCs using three different strategies. This allows for comparison of both the results and the methods for obtaining them. The variability of the obtained results may influence our confidence in the estimated distribution of hydrological properties. Another reason to do this is to evaluate the advantages and drawbacks of various approaches, which may differ given the objective of a study, the number of parameters to estimate, available computing facilities, and details of the forward model. Below, to evaluate model misfit we used the root-mean-square (RMS) difference between simulated and measured BTCs using all available points in time, which have a relatively constant sampling time interval. Although we tested other misfit evaluation techniques (first and second moments, quartiles, and first arrival time with areas below the BTCs), they did not significantly change the overall results.

5.1. Grid Search Strategy

[22] We first used a grid search strategy to evaluate the main properties involved in the transport process and to perform sensitivity analysis. The grid search approach is based on the evaluation of one or more parameters through a range of values, with the possibility of starting with a large-scale search grid and subsequently refining it. This means the selected search domain, which is defined by a range of values that we assign to each parameter, is sampled fully for each dimension of the problem without optimization, so this is the most intensive computing method.

[23] The main advantages of the grid search approach are that (1) the process is easily parallelized, even on a network of independent computers (or distributed memory cluster) as no optimization process is performed; (2) the objective function can be evaluated after the simulation process has been completed, which in turn, allows for evaluation of multiple objective functions; (3) this approach does not require optimization, regularization, or linearization, and thus is not sensitive to convergence problems that

may be significant in large, complex models; and (4) the sensitivity of each parameter can be evaluated because the domain of the solution is directly sampled. However, a simple visual observation of the solution domain is only possible with a very small number of parameters. Finally, we note that the computing efficiency of this approach can be improved using a Monte Carlo strategy to avoid visiting all the cells of the grid, or a multistep approach where only a small number of parameters are evaluated sequentially and/or the grid is iteratively redefined [e.g., *Stein*, 1987; *Sambridge and Mosegaard*, 2002].

5.2. Simulation of Hydrological Property Distributions

[24] Stochastic simulations are used for a number of purposes in earth sciences including heterogeneity reconstruction, uncertainty assessment, and optimization [e.g., *Deutsch and Journel*, 1998; *de Marsily et al.*, 2005]. Such simulations can be unconditional or conditioned to available data at some locations. Although numerous approaches exist [e.g., *Deutsch and Journel*, 1998; *Christakos*, 2000; *Tarantola*, 2005], they are similar in that they require some statistical information like correlation functions, mean and variance, or some data to constrain these values. Here we use the well-known sequential Gaussian simulation (SGS) algorithm from the Geostatistical Software Library (GSLIB) [*Deutsch and Journel*, 1998], and we use information available at the BHRS to define statistical parameters in the simulation process. We use SGS to generate a large number of unconditional simulations of hydrological property distributions (K and ϕ, which are assumed to be positively correlated and linearly related) and to investigate the commonalities among those realizations, which reproduce the measured BTCs reasonably well. For clarity, we note that simulations used here to find models that fit the BTCs are not conditioned to any data point, whereas simulations conditioned to measured hydrological data are more commonly performed to assess the uncertainty in predicting hydrological processes such as flow and/or transport.

[25] In particular, this procedure generates and supports investigation of heterogeneity at a smaller scale than achievable with other parameter estimation approaches. Generation of small-scale heterogeneity of hydrological properties may support (1) improving fits to BTCs and (2) finding models showing the nonuniqueness in the solution domain (i.e., finding models that fit the data to a similar level and evaluating how different they are). The only critical issue related to this approach is the need to reduce computing time by constraining the simulation process with prior statistical information, such as the global mean and variance of the property distributions, which limit the search domain and thus may influence the results.

5.3. Regularized Least Square Inversion

[26] Approaches for the inversion of hydrological flow and/or solute transport consist of sequentially updating the parameters in a model to minimize the difference between the calculated and observed data. Reviews of the state of the art of such techniques are numerous [e.g., *Menke*, 1984; *Yeh*, 1986; *de Marsily et al.*, 1999; *Aster et al.*, 2005; *Carrera et al.*, 2005; *Tarantola*, 2005]. Here we apply a deterministic inversion that optimizes the value of parameters at pilot points [e.g., *Doherty*, 2003; *Carrera et al.*, 2005; *Johnson et al.*, 2009] subject to an objective function that contains a least squares data misfit criterion and a regularization term. The objective function is

$$\Phi(m_{\text{est}}) = \|W_c(G_c[m_{\text{est}}] - C_{\text{obs}})\|^2 + \beta\|W_m m_{\text{est}}\|^2, \quad (1)$$

where $G_c(m_{\text{est}})$ and C_{obs} are vectors of length n, where n is the number of measurements. $G_c(m_{\text{est}})$ contains the concentration data predicted from modeling the solute transport (using SEAWAT) given the estimated parameters m_{est} and C_{obs} as the measured concentration data. W_c is the data weighting matrix of size $n \times n$, where each diagonal element is the reciprocal of the estimated standard deviation of each measurement and the off-diagonal elements are zero if data errors are uncorrelated. The second term of equation (1) is the regularization term, where W_m (of size $m \times m$, where m is the number of unknowns) allows directional smoothness constraints and β is a scalar weighting parameter that defines how much the regularization influences the process of decreasing the misfit between the calculated and measured data; W_m and β can be directionally dependent.

[27] At each iteration, we update parameters by finding a step Δm_{est} that reduces the value of the objective function. After developing equation (1) through a Taylor series approximation, the final equation is

$$\begin{aligned}(J_c^T W_c^T W_c J_c + \beta W_m^T W_m)\Delta m_{\text{est}} \\ = J_c^T W_c^T W_c(C_{\text{obs}} - G_c[m_{\text{est}}]) - \beta W_m^T W_m m_{\text{est}},\end{aligned} \quad (2)$$

where the concentration Jacobian matrix J_c is obtained using a finite difference approach [*Johnson et al.*, 2009]. The perturbation amount Δm_{est} is solved for each iteration using a conjugate gradient least squares algorithm. After each iteration, the new estimated m_{est} is obtained by adding to the previous solution the updated vector Δm_{est}, which is rescaled on the basis of finding the optimum scaling that produces the greatest decrease to the objective function. The ideal number of processors to solve such a problem is one plus the number of pilot points for which the sensitivity matrix needs to be evaluated. In this study we used an eight-core cluster and the parallel computation tools available in MATLAB.

[28] Also, in this study, as in hydrological inversion in general, the vector m_{est} contains the estimated hydraulic parameters at a set of pilot points only. These values distributed at a few cells in the model domain are then used to generate values at each grid element. Generally, this is done through interpolation, kriging, and/or simulation processes [e.g., *Certes and Demarsily*, 1991; *Ramarao et al.*, 1995; *Rubin et al.*, 2010]. The choice of the appropriate method depends on the number of pilot points and some prior idea of the required heterogeneity to fit the measured data, which can be gained by preliminary modeling and data mining, and/or from structural information available from other geophysical or hydrological data at the site. The more pilot points that are used to characterize a spatial parameter distribution, the less important is the choice of interpolation method [*Doherty*, 2003]. Also, it is important to note that kriging the values from the pilot points,

although producing some smoothness in the model, does not control the correlation between the pilot points. In fact, the overall variability of the parameters defined at the pilot points is controlled by the regularization term only.

[29] Although the implementation of this regularized inversion approach looks more complex than previously discussed methods, once the equation system is set, advantages include (1) the significant reduction of computing resources (i.e., although it is less easy to parallelize on a network of independent computers), (2) the ability to find a model of minimal variability that reproduces the data, and (3) the ability to solve for a relatively large number of parameters in general. In particular, the reduction of computing time allows us to more easily evaluate various parameterizations and/or problem configurations and their effect on the resulting distribution of hydrological properties.

6. Results

6.1. Parameter Estimation Through a Grid Search Strategy

[30] The tracer test data are first investigated with a grid search strategy to evaluate some simple model configurations (i.e., defined by a relatively small number of parameters) that may allow us to reproduce these data to a first order. We first evaluate a homogeneous model, which, because a limited number of parameters are evaluated, is relatively uniformly sampled and the residual misfit between predicted and measured concentration data can be observed throughout the full considered parameter domain. Then we investigate how a relatively simple layered model can fit measured data and better predict measured concentrations than a homogeneous model.

6.1.1. Homogeneous Model

[31] Here we evaluate some broad ranges of ϕ and K for a homogeneous model. Figure 7 shows RMS residuals between all predicted and measured concentrations for each set of evaluated properties. It is clear that relatively similar fits result from correlated changes in ϕ and K. This is explained by the link between ϕ and K in the transport equations. However, the RMS residual increases significantly for ϕ lower than ~0.28, so it appears that the coupling between ϕ and K has some influence on the sinking (density effect) of the tracer, and the ϕ and K values need to be in a limited range to fit the BTCs acceptably. It is important to note that only enforcing very high K values at some depth interval does not allow draining of the tracer and improved fitting to the BTCs. Instead, both high ϕ and K seem to be necessary in order for a larger amount/mass of tracer to sink rapidly such that adequate fits to BTCs are obtained.

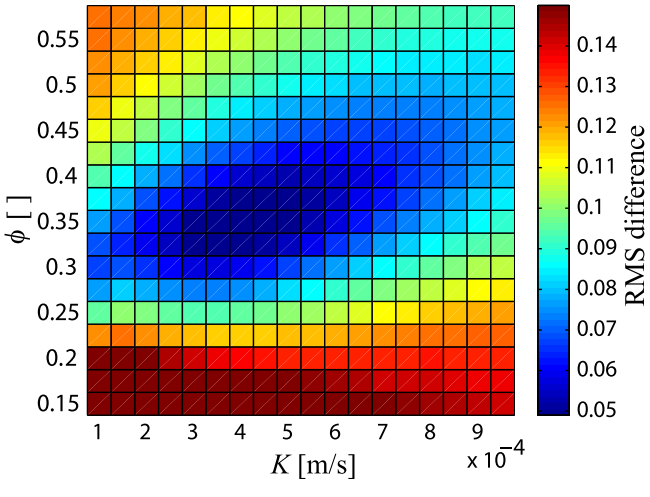

Figure 7. Grid search results for the evaluation of homogeneous models of ϕ and K that support the prediction of solute concentration data and are considered here in regard to the root-mean-square (RMS) difference between predicted and measured BTCs at well A1.

[32] Figure 8 shows the predicted concentration from the ϕ and K model with the best fit of the BTCs. Although this model gives acceptable fits to first arrival time and some variations in the BTCs, and thus is worthy for a first-order evaluation of hydrological data and properties, more heterogeneity is needed to get better fits at each measurement depth level. For example, the velocity in the bottom part of the model should be larger to better fit the early arrival time between 836.7 m and 837.45 m. Also, the homogeneous model does not allow the tracer to sink enough and thus predicted concentrations between 838.45 m and 839.95 m are too high and the breakthrough arrival time is too early there.

6.1.2. Layered Model

[33] A grid search approach is used to evaluate a three-layered model with various possible layer thicknesses and K and ϕ values. For each layer, the ranges of values evaluated for these parameters were defined using prior information from a GPR velocity model [*Dafflon et al.*, 2011] and ϕ logs [*Barrash and Clemo*, 2002], and from the grid search using a homogeneous model. Thus, ϕ, K, and the top of the layer are evaluated for the interval [0.15; 0.25; 0.35], [0.5e-4; 1e-4; 3e-4], and [850], respectively, for the upper layer; [0.25; 0.35; 0.45], [2e-4; 4e-4; 6e-4], and [838.475; 838.075; 837.675], respectively, for the central layer; and [0.25; 0.35; 0.45], [2e-4; 4.5e-4; 7e-4], and [837.275; 836.675; 836.075], respectively, for the lower layer.

[34] The reason for evaluating a three-layered model and such a small number of parameters and range of values is the significant demand on computing resources, which in this case was ~20 dual-core computers running at night for 1 week to perform 6561 simulations. Computing time would increase exponentially with the number of evaluated parameters, such as, for example, evaluating various dispersivity values. Hydrologically, use of a three-layered model rather than a five-layered model for the BHRS (e.g., *Barrash and Clemo* [2002]; and see Figures 2 and 3 in this work) is justified because flow and transport during the tracer test occurred primarily or exclusively in the middle layers of the system.

[35] Figure 9 shows the model with the lowest RMS difference (0.0421) between predicted and measured concentrations among all of the evaluated models. The fit to BTCs is improved compared to Figure 8, mainly because K and ϕ are now larger below 838.075 m (also see Figure 2). In fact, the ratio of K to ϕ is higher which implies higher velocity and earlier arrival of large concentrations below 837.45 m. The arrival times and principal behaviors of the BTCs are relatively well reproduced. The concentrations

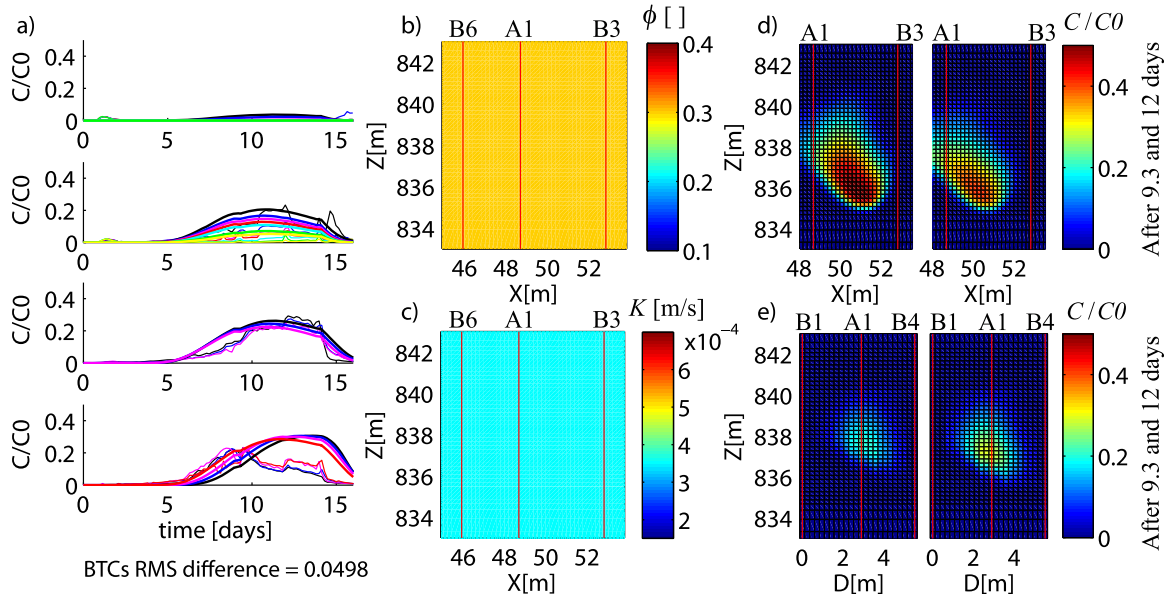

Figure 8. Best fit model obtained through a grid search evaluation of homogeneous K and ϕ fields: (a) measured (thin lines) and calculated (thick lines) BTCs in 20 depth intervals along A1 (shown in Figure 5), (b) estimated ϕ, and (c) estimated K distributions, shown along profile B6–B3; (d–e) simulated concentration after 9.3 and 12 d observed along profiles A1–B3 and B1–A1–B4, respectively.

simulated between B1–B4 (Figure 9e) show the presence of the tracer at depths that are similar to those in time-lapse GPR images (Figure 6). Also, the significant changes in ϕ and K at 838.075 m are consistent with the sharp boundary between units 2a and 2b at 838 m in the GPR tomogram along profile B6–A1–B3 (Figure 3) and in the ϕ logs where high-porosity lenses >0.30 occur in, and likely between, wells B3 and A1 between 838 and 836 m elevation. Indeed, the results indicate that a simple three-layer model can explain a large part of the information contained in the data, and thus that principal behaviors in the BTCs can be related to hydrostratigraphic layers or large hydrological structures. However, some behaviors in the measured BTCs that are still not well reproduced by the model, such as the

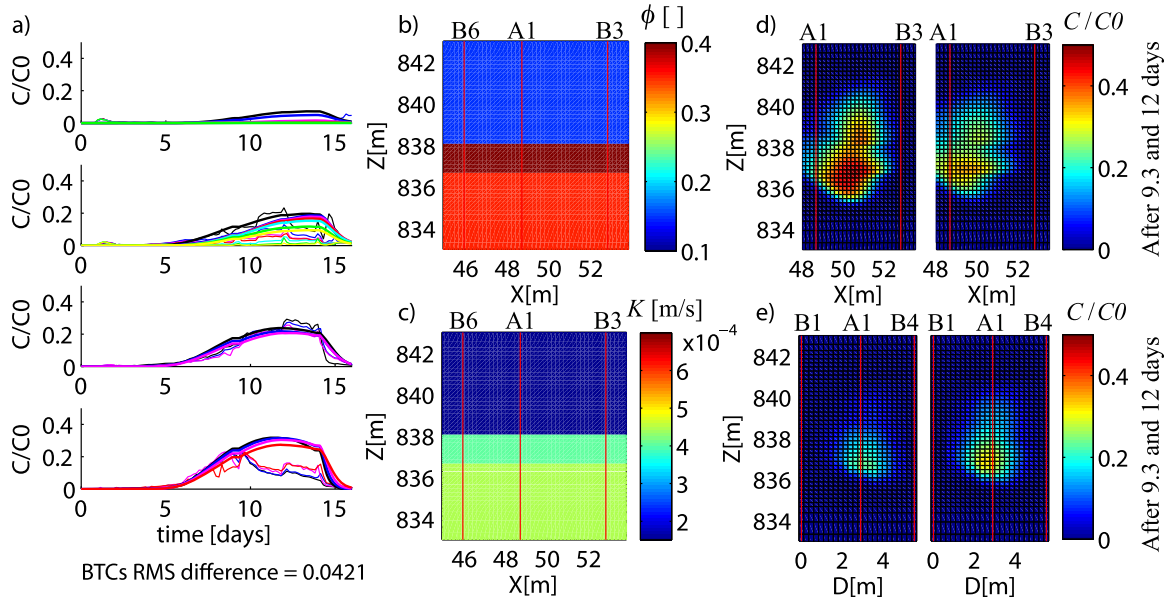

Figure 9. Best fit model obtained through a grid-search evaluation of a three-layer model with moveable boundaries: (a) measured (thin lines) and calculated (thick lines) BTCs in 20 depth intervals along A1 (shown in Figure 5), (b) estimated ϕ, and (c) estimated K distributions, shown along profile B6–B3; (d–e) simulated concentration after 9.3 and 12 d observed along profiles A1–B3 and B1–A1–B4, respectively.

decrease in concentration after ~10 d in the BTCs measured between 836.7 and 837.45 m, will be investigated next by simulating small-scale variability in the hydrological models.

6.2. Stochastic Simulation of Hydrological Properties

[36] We perform stochastic simulations to obtain realizations of K and linearly related ϕ with small-scale variability that may improve matches to the measured BTCs. We investigate this with two approaches: (1) with the three-layer model shown in Figure 9 (i.e., an a priori setting of the large-scale structures) to examine improvement in fitting the BTCs due to the simulation of small-scale heterogeneity in each layer, and (2) with simulations that reproduce small- and large-scale structures without prior layering specified in the model to evaluate the extent to which such simulations show similarities when reproducing the measured data to some degree. It is important to note that parameterizations of these two simulation approaches, involving means and covariance functions, are different due to the different scale at which heterogeneity needs to be defined in each.

6.2.1. Unconditional Simulations in a Three-Layer Model

[37] From the three-layer models with the best fit BTCs from the grid search estimation process (section 6.1.2), an arbitrary number of 2500 simulations were performed in each layer to create small-scale heterogeneity. The mean and variance of the simulations in each layer were defined based on the results obtained earlier through the grid search strategy (Figure 9). For the bottom, middle, and top layers, mean K is set to 5.3e-4, 4.2e-4, and 2e-4 m s^{-1}, respectively (i.e., similar to values from fully penetrating pumping tests at the BHRS [*Barrash et al.*, 2006]), with standard deviations of 2e-4, 2e-4, and 1e-4, respectively, and the minimum value is set to 0.5e-4 m s^{-1}. Also, ϕ is linked to K with a linear relationship, which is consistent with the positive correlation observed between ϕ and K in the three-layer model, and with the need to set some parameters using prior results to simplify the problem. For the bottom, middle, and top layers, the resulting mean ϕ values are 0.33, 0.33, and 0.2, respectively, with standard deviations of 0.08, 0.08, and 0.08, respectively, and the minimum value is set to 0.05. The correlation lengths were defined based on the vertical correlation length of 1.2 m obtained from fitting the correlation function from the ϕ logs for the various depth intervals, which in turn, is similar to values determined by *Barrash and Clemo* [2002]. Other values could be used to parameterize the simulations, but these are reasonable choices and the purpose here is more to evaluate improvement in the BTCs by introducing small-scale structures than to estimate all possible solutions.

[38] Figure 10 shows the hydrological realization with the best fit to the measured BTCs (RMS residual is 0.0248). This simulation shows the extent to which small-scale structure can improve fits to the BTCs, including reproduction of some small-scale behaviors in the BTCs. The five "best" simulations have residuals between 0.0248 and 0.0287 and the RMS residual distribution of all the simulations has a Gaussian behavior that can be fit with mean and standard deviations of 0.047 and 0.008, respectively. This shows that a relatively small number of solutions significantly improve the three-layer model (Figure 9) and that the simulation of such small-scale structures can strongly influence behaviors in the predicted concentration curves (Figure 10). Also, we note that high ϕ values were required to improve the fits, which again, are consistent with the local structure of high ϕ lenses below 838 m elevation (i.e., especially between 838 and 836 m but also between 836 and

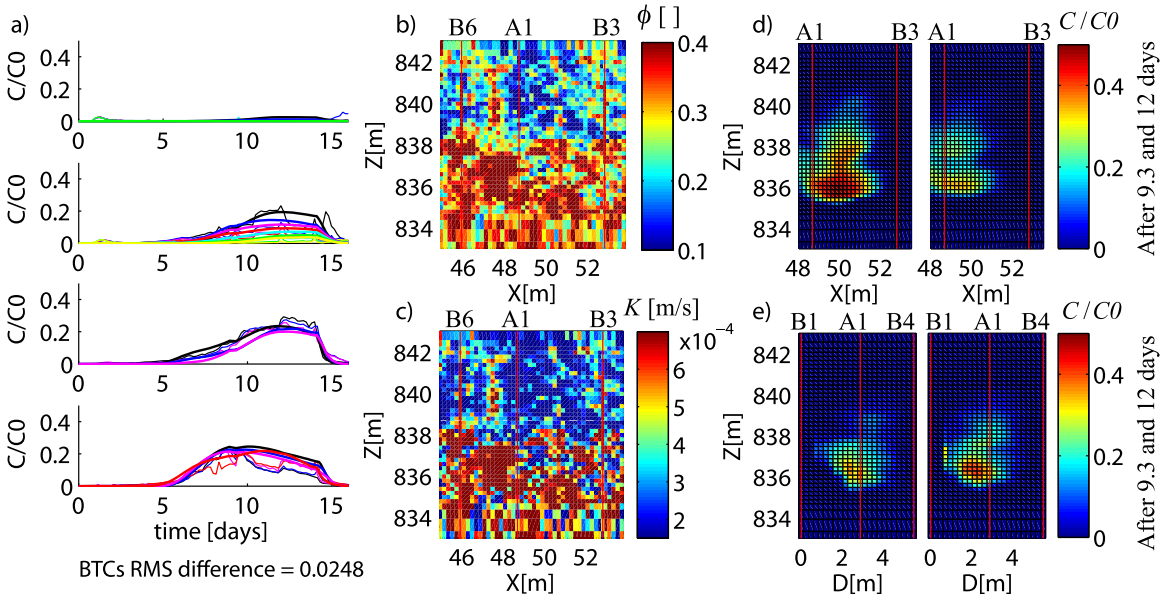

Figure 10. Best fit model obtained through a simulation approach using the geometry of the three-layer model shown in Figure 9: (a) measured (thin lines) and calculated (thick lines) BTCs in 20 depth intervals along A1 (shown in Figure 5), (b) simulated ϕ, and (c) simulated K distributions, shown along profile B6–B3; (d–e) simulated concentration after 9.3 and 12 d observed along profiles A1–B3 and B1–A1–B4, respectively.

834 m; see Figure 2) and corroborates the results of the grid search approach. Finally, although the model shown in Figure 10 gives a very good fit to the BTCs overall, some features are still not well reproduced, especially some high-frequency variations in the BTCs, such as, for example, the decrease in concentration after ~10 d and then the increase again after ~12 d in the BTCs measured between 836.7 and 837.45 m elevation (Figure 10). In fact, incomplete fitting of BTCs is common [e.g., *Hyndman et al.*, 2000; *Scheibe and Chien*, 2003; *Johnson et al.*, 2009] and it is difficult to predict the degree to which real BTCs can be reproduced. This is mainly related to limitations in the understanding and modeling of transport processes, boundary effects, and the parameters of influence.

6.2.2. Unconditional Simulations in a One-Layer Model

[39] Now we investigate if structural similarities can be found between realizations that fit the BTCs relatively well but have not been conditioned with prior layer information. To this end, simulations of K and linearly related ϕ fields are performed based on previously estimated global statistics at the BHRS. The global means of K and ϕ for the simulations are set to 4e-4 m s^{-1} and 0.3, respectively, and their standard deviations are set to 2e-4 and 0.1, respectively. These are similar to thickness-averaged K values from fully penetrating pumping tests [*Barrash et al.*, 2006] and important ϕ structures at B3 and A1 (Figure 2). Minimum ϕ and K values are set to 0.05 and 0.5e-4, respectively. Horizontal and vertical correlation lengths have been set to 8 m and 1.6 m, respectively. These lengths are based on the previous results including structural information from GPR data (Figure 3), and are similar to lengths obtained from the geostatistical analysis of K from slug test results at the BHRS [*Cardiff et al.*, 2011].

[40] The model with the lowest RMS difference between predicted and measured BTCs is shown in Figure 11, and Figure 12 shows four other realizations with next best fits to the BTCs. These five models have RMS residuals between 0.026 and 0.0276 and show structural similarities with (1) layers of higher K and ϕ between 836 and 838 m (Figures 11, 12c, and 2), which also are visible in the mean of these simulations (Figure 12b); and (2) the strong change to lower ϕ and K above 838 m (that also occurred in the three-layer model; see Figures 9 and 10) compared to the strong change in GPR velocity and ϕ at the same elevation (Figures 2 and 3a). Compared to the three-layer model obtained by grid search (Figure 9), the BTCs are better fit with the lowest RMS models (Figures 11 and 12c), which is due to more structural variability in the simulations. In particular, the choice of a relatively large lateral correlation length compared to the size of the model allows some variation in the mean or variance locally, which in turn, allows a large distribution of RMS differences (Figure 12a) and improves the possibility of finding simulations with particular features that are able to fit the data relatively well (Figures 11 and 12c), even if such simulations are few.

6.3. Parameter Estimation Through a Regularized Inversion

[41] In section 6.3 we perform regularized inversion with ϕ as (1) a predefined constant value (section 6.3.1), (2) an estimated parameter in the inversion process (section 6.3.2), and (3) a predefined distribution obtained from 3-D ϕ reconstruction from geophysical data (section 6.3.3). Such

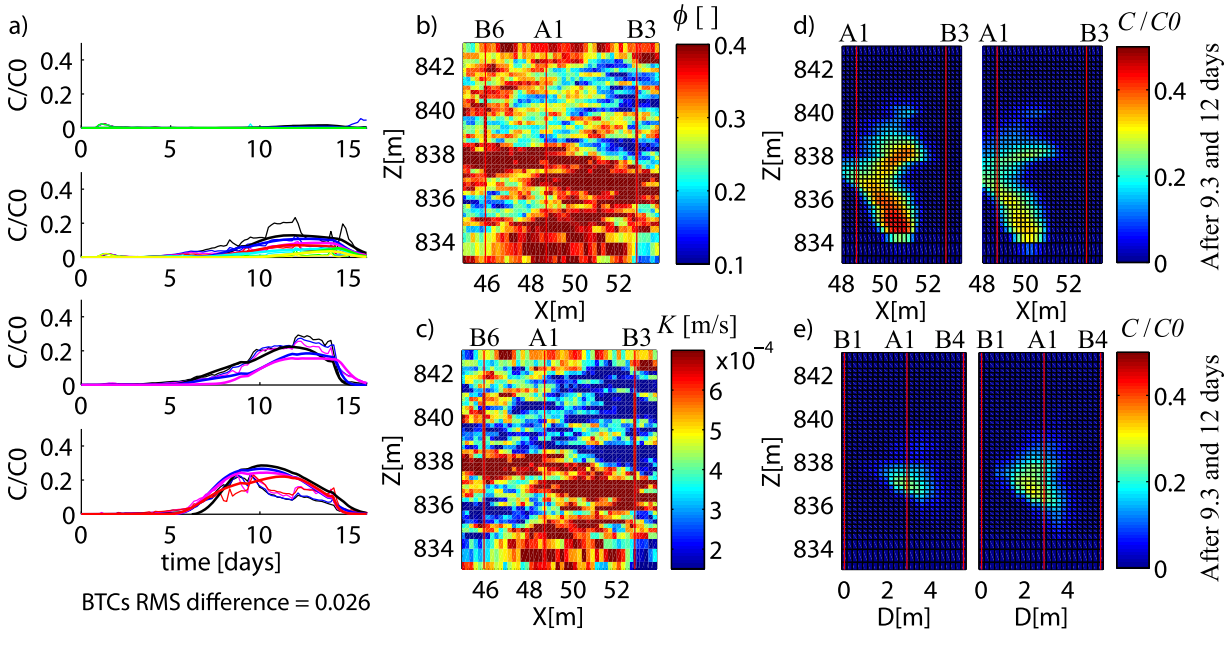

Figure 11. Best fit model obtained through a simulation approach without layers: (a) measured (thin lines) and calculated (thick lines) BTCs in 20 depth intervals along A1 (shown in Figure 5), (b) simulated ϕ, and (c) simulated K distributions, shown along profile B6–B3; (d–e) simulated concentration after 9.3 and 12 d observed along profiles A1–B3 and B1–A1–B4, respectively.

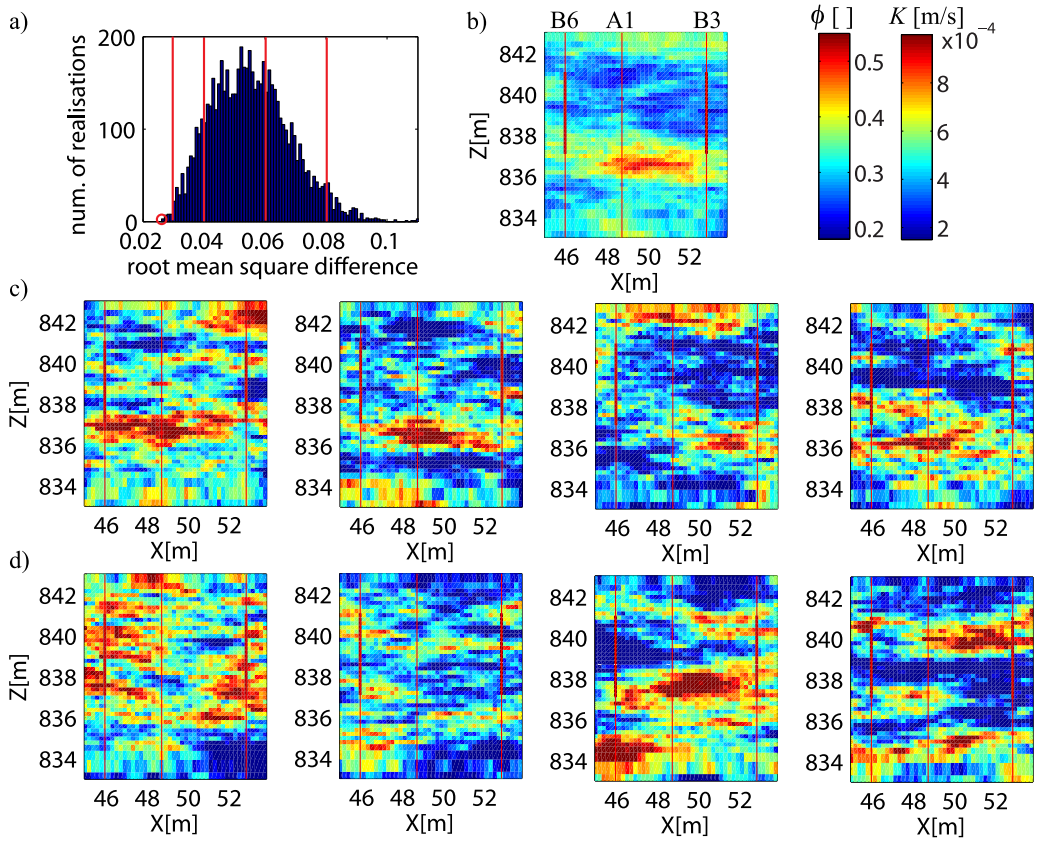

Figure 12. (a) Distribution of the root-mean-square difference between measured and predicted concentrations obtained from 6000 simulations of the hydrological structure. (b) Mean of the linearly related K and ϕ realizations of the five best fits to the measured BTCs (red circle in Figure 12a). The best-fitting realization is shown in Figure 11. (c) The next four best fits. (d) Four realizations (red lines in Figure 12a) of various fits of the measured concentration to show the variability generated through the used simulation approach.

regularized inversion allows us to evaluate (1) the role of ϕ in the transport process, (2) whether it is necessary to include ϕ variation in the estimation procedure, (3) the value of using geophysical data to develop the ϕ model, and (4) the consequences of choices made in parameterizing the inversion process on the results and on fitting the measurements. All of the inversions presented here were performed using seven pilot points positioned at a succession of depths at a given x–y coordinate position (shown with black dots in Figures 13–18). A layered model is obtained by kriging with very large lateral correlation lengths. These choices represent a trade-off between simplifying the problem and increasing the number of pilot points and problem dimensions, where choosing the more complicated approach would (1) significantly increase computing demand, (2) require more information to constrain/control an increasingly ill-posed problem, and (3) risk estimation of a model with more heterogeneity than is needed to explain the data. Indeed, the large hydrological units visible at the BHRS from other sources of information (Figure 3) indicate that a simplified approach might estimate K and ϕ structures sufficiently well to match the BTCs to a first order. The regularization for the inversion process was imposed on the hydrological model using second-derivative smoothness constraints in the x, y, and z directions. Note too, that we also tested a Bayesian approach which yielded similar results using an estimated prior mean and covariance matrix.

6.3.1. Inversion of Hydraulic Conductivity Using a Predefined Constant Effective Porosity

[42] It is quite common to set ϕ to a constant value in an inversion procedure to estimate K because the influence of the ϕ distribution on the flow and transport process is often assumed to be small compared to that of K. Doing this also reduces computing demand. Here we consider two cases where ϕ is assumed known and (1) equal to 0.3, and then (2) equal to 0.2. Both values can be considered as reasonable estimates for the aquifer at the BHRS based on the available measurements at the site.

[43] Figure 13 shows the resulting K and the predicted BTCs when ϕ is set to 0.3. The RMS difference between the predicted and measured BTCs is 0.0305, and thus can be considered as relatively good. Also, the main depth variations in K are relatively consistent with some previous results (Figures 9–11) and with high-porosity structures seen in ϕ logs (Figure 2) below 838 m between wells B3 and A1 and imaged by geophysical data (Figure 3). However, when ϕ is 0.2 (Figure 14), the resulting K model

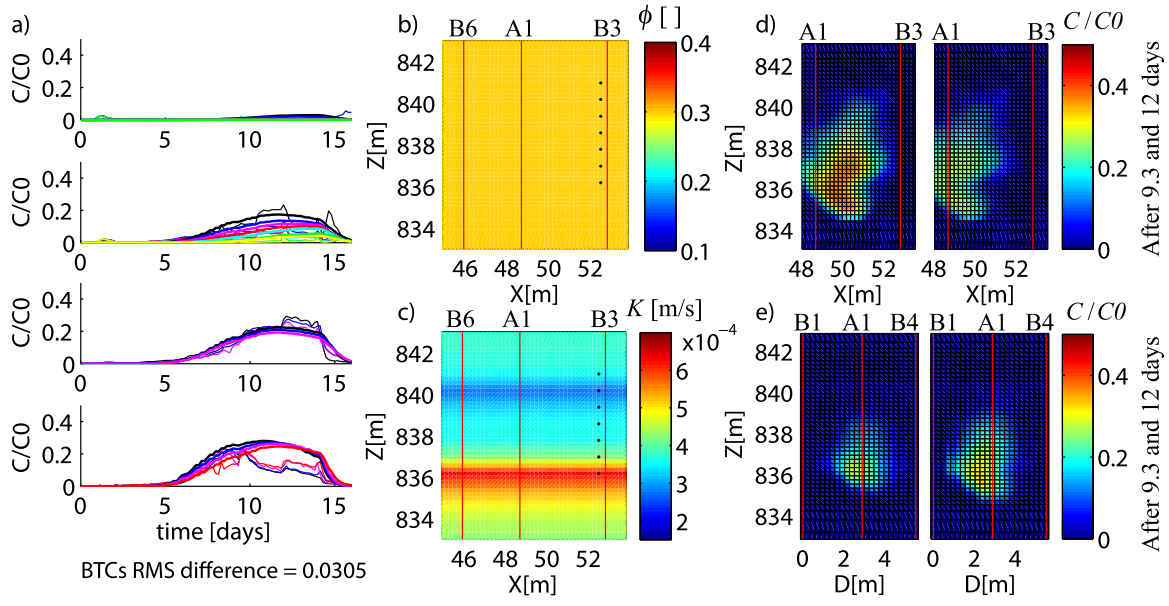

Figure 13. Best fit model obtained through a regularized least square inversion to estimate K distribution: (a) measured (thin lines) and calculated (thick lines) BTCs in 20 depth intervals along A1 (shown in Figure 5), (b) assumed known ϕ, and (c) estimated K distributions, shown along profile B6–B3; (d–e) simulated concentration after 9.3 and 12 d observed along profiles A1–B3 and B1–A1–B4, respectively.

yields a poorer RMS difference (0.0484) between predicted and measured BTCs. Also, the simulated plume crossing profile B1–B4 at various times (Figure 14) is not qualitatively similar to the time-lapse GPR tomography (Figure 6). That is, a difference of 0.1 in the assumed known ϕ has a large influence on the results, and a ϕ value that is too low, in this case 0.2, does not allow the simulated plume to sink enough regardless of the associated K distribution estimated through the inversion process. Importantly, these results are consistent with those obtained from the grid search approach applied to a homogeneous hydrological model, where low ϕ does not allow good fits to the measured BTCs (Figure 7). Finally, these results show that setting ϕ to predefined values, which is quite common in hydrological studies, can significantly influence the results from the inversion process.

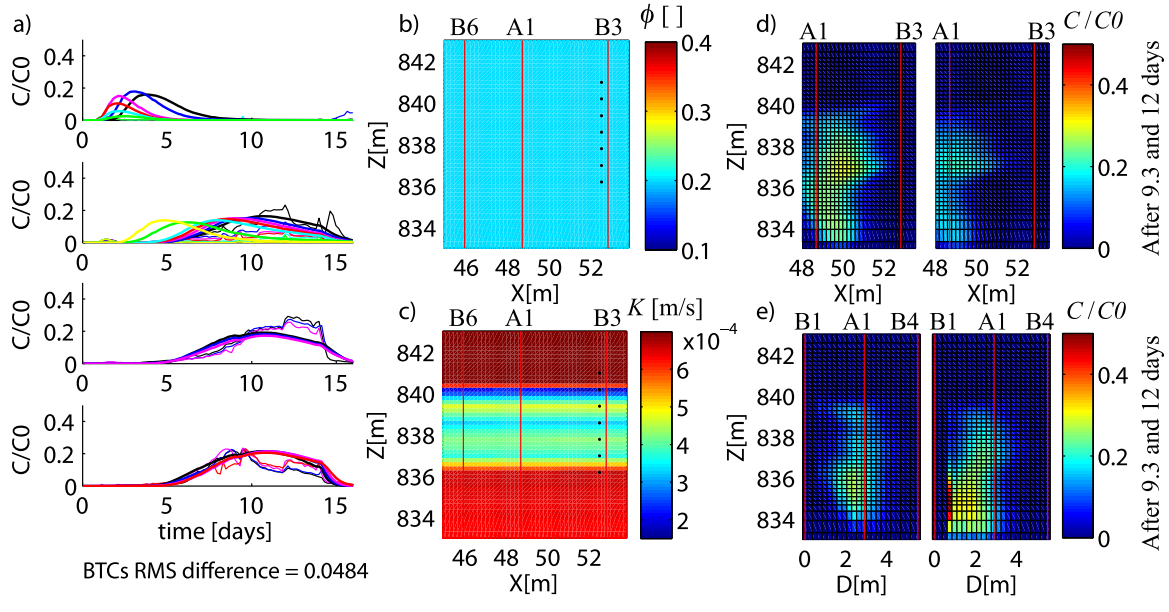

Figure 14. Best fit model obtained through a regularized least square inversion to estimate K distribution: (a) measured (thin lines) and calculated (thick lines) BTCs in 20 depth intervals along A1 (shown in Figure 5), (b) assumed known ϕ, and (c) estimated K distributions, shown along profile B6–B3; (d–e) simulated concentration after 9.3 and 12 d observed along profiles A1–B3 and B1–A1–B4, respectively.

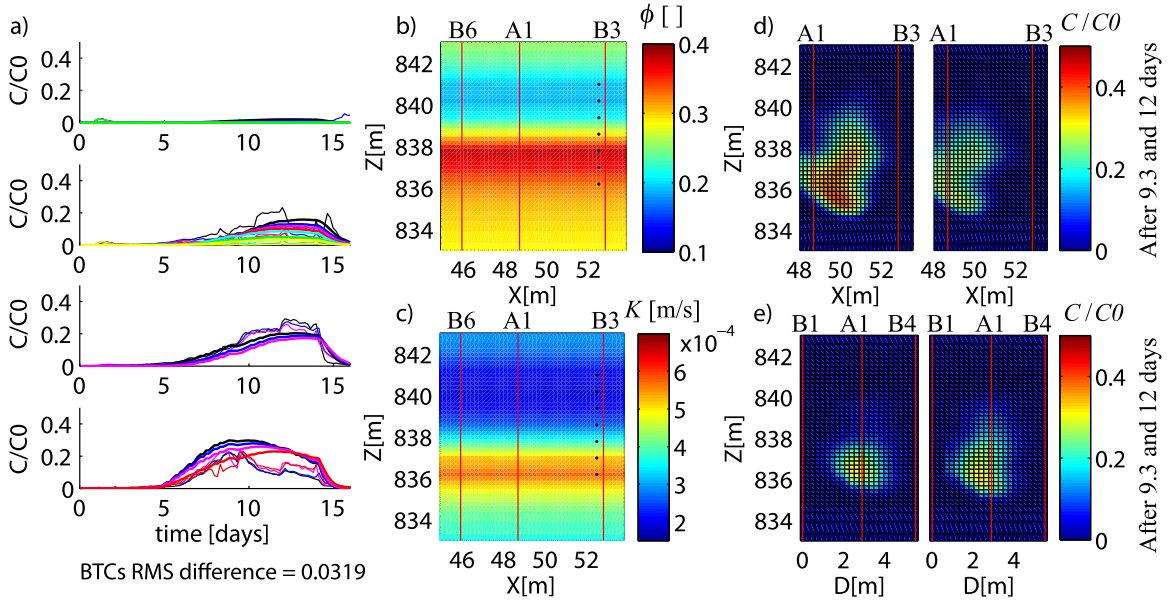

Figure 15. Best fit model obtained through a regularized least square inversion to estimate K and ϕ distributions and using a relatively large regularization weighting coefficient: (a) measured (thin lines) and calculated (thick lines) BTCs in 20 depth intervals along A1 (shown in Figure 5), (b) estimated ϕ, and (c) estimated K distributions, shown along profile B6–B3; (d–e) simulated concentration after 9.3 and 12 d observed along profiles A1–B3 and B1–A1–B4, respectively.

6.3.2. Inversion of Hydraulic Conductivity and Effective Porosity

[44] Deterministic inversion has been performed to estimate both K and ϕ for the following two cases. First (Figure 15), the regularization weighting coefficient β is selected based on the trade-off between data fit and model structure. This means the structure, or more exactly the values at the pilot points, are enforced for continuity (smoothness). Second (Figure 16), a smaller regularization coefficient is applied and thus the variability between the estimated values at the pilot points may be higher. Note that although kriging controls some spatial variability between the pilot

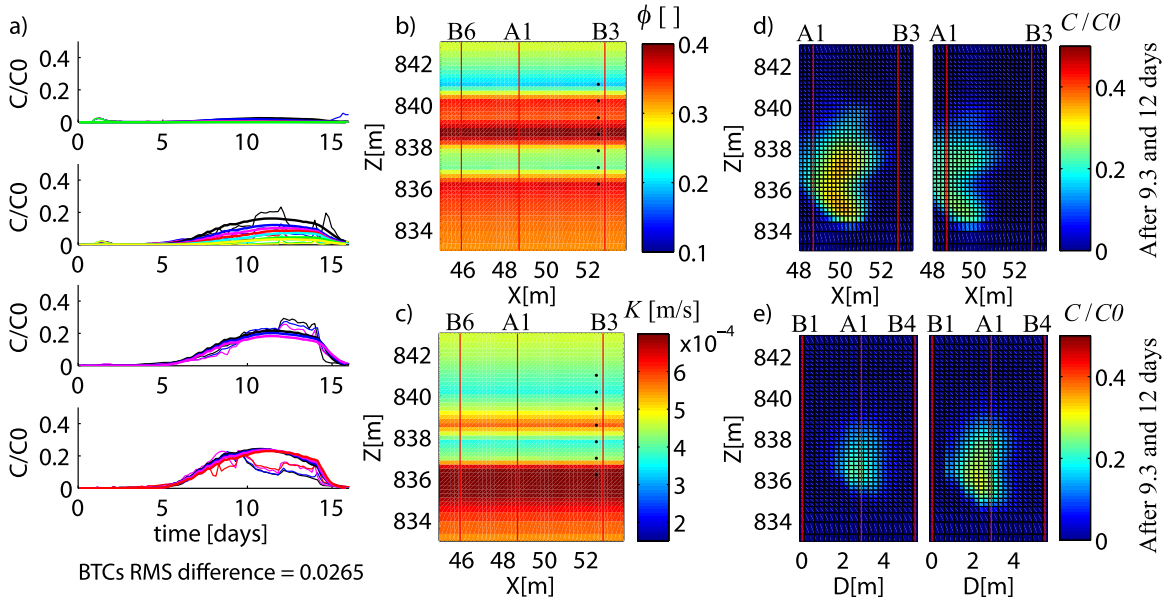

Figure 16. Best fit model obtained through a regularized least square inversion to estimate K and ϕ distributions and using a relatively low regularization weighting coefficient: (a) measured (thin lines) and calculated (thick lines) BTCs in 20 depth intervals along A1 (shown in Figure 5), (b) estimated ϕ, and (c) estimated K distributions, shown along profile B6–B3; (d–e) simulated concentration after 9.3 and 12 d observed along profiles A1–B3 and B1–A1–B4, respectively.

points, the variability between the values estimated at the pilot points is controlled only by the regularization.

[45] For the first case (Figure 15), the RMS difference (0.0319) between the estimated and measured BTCs is relatively good and the K and ϕ distributions show nice agreement with previous results (Figures 9, 10, 11, and 13) and geophysical data (Figure 3) and ϕ logs (Figure 2), especially with regard to increases with depth at ~838 m. However, when inverting the data with a small regularization weighting parameter (Figure 16), the estimated K and ϕ distributions are very different. In contrast to Figure 15, K and ϕ distributions in Figure 16 show strong depth-related fluctuation and are negatively correlated. Here the low regularization parameter has enabled more variability between the values estimated at the pilot points, which has led to a completely different solution with a lower RMS difference (0.0265). However, these results are not very realistic considering the poor similarity in variability between the estimated ϕ (Figure 16) and the ϕ log (Figure 2) and the geophysical data (Figures 3 and 6). That is, the choice of the regularization parameter may result in very different solutions even when structures are evaluated at a relatively large scale. This nonuniqueness is likely driven by the fact that both K and ϕ distributions are estimated through the inversion process, and more so without any assumption about a relation between them. So this again shows the advantages of and need to constrain the estimation of hydrological properties although, at the same time, constraints applied on the inversion process can strongly influence the result and thus need to be reliable.

6.3.3. Inversion of Hydraulic Conductivity Using a Predefined 3-D Porosity Distribution

[46] From the above results it is clear that estimation of the 3-D ϕ distribution can be important to constrain the inversion of hydrological data to estimate K. Here this is done by using a ϕ realization obtained at the BHRS from ϕ logs (e.g., Figure 2) and GPR cross-hole velocity tomograms with a simulated-annealing approach (e.g., Figure 3b). The ϕ realization used here can be considered as representative of a large number of realizations in the context of this study.

[47] Figure 17 shows the K distribution estimated through the regularized inversion when a simulation of ϕ from geophysical data is used to define the ϕ model. However, the estimated model of K does not provide a satisfactory RMS difference (0.0759) between the predicted and measured BTCs. In fact, this result is similar to that obtained when setting the ϕ equal to 0.2 prior to the inversion (Figure 9) and is similar to the results from the grid search approach applied to a homogeneous model (Figure 7). In all these cases, having ϕ values that are too low results in models that poorly match the measured BTCs. Given these results we decided to rerun the same deterministic inversion after adding 0.1 to the estimated ϕ distribution based on the previous grid search and inversion results. In this case (Figure 18), the model of K is more realistic and results in a much smaller RMS difference (0.0308) between predicted and measured BTCs. Furthermore, the estimated K distribution now agrees well with previous results showing similar features (Figures 9, 10, 13, and 15) and shows a better agreement with the GPR time-lapse plume images (Figure 6) compared to Figure 17. This result shows that the structure estimated from geophysical data can help constrain the solution domain and is not meant to indicate that estimated ϕ from neutron logs or other geophysical data will necessarily need to be corrected for use in the estimation of K. In fact, several factors may explain this including (1) the difficulty of modeling and understanding the near-well processes, especially density-effects during the injection and early

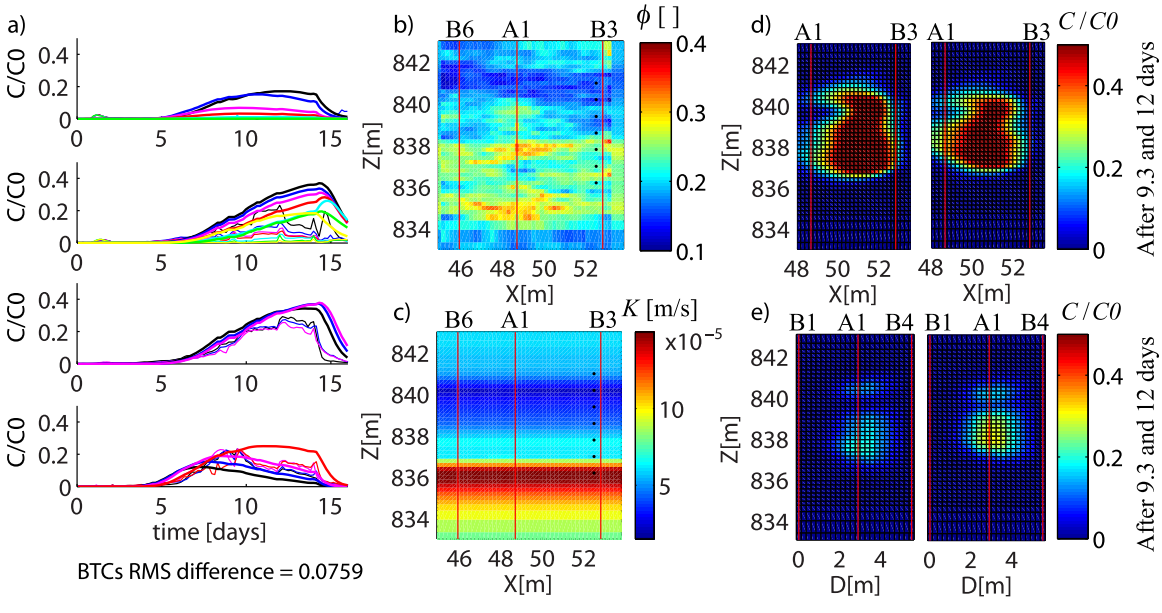

Figure 17. Best fit model obtained through a regularized least square inversion to estimate K distribution: (a) measured (thin lines) and calculated (thick lines) BTCs in 20 depth intervals along A1 (shown in Figure 5), (b) ϕ from geophysical and well data, and (c) estimated K distributions, shown along profile B6–B3; (d–e) simulated concentration after 9.3 and 12 d observed along profiles A1–B3 and B1–A1–B4, respectively.

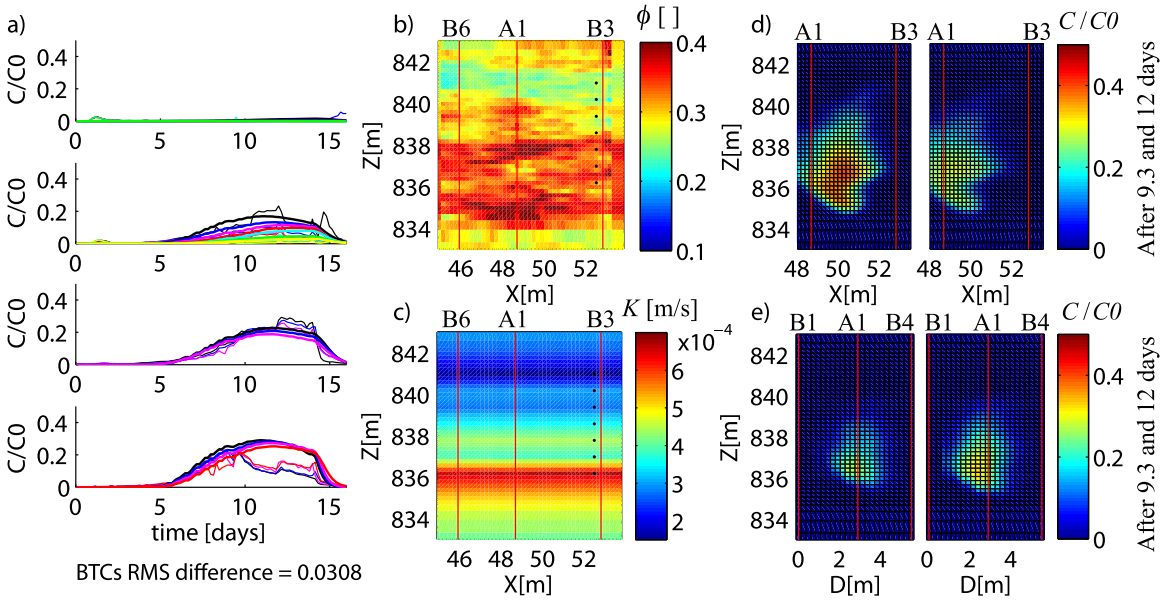

Figure 18. Best fit model obtained through a regularized least square inversion to estimate K distribution: (a) measured (thin lines) and calculated (thick lines) BTCs in 20 depth intervals along A1 (shown in Figure 5), (b) ϕ from geophysical and well data with 0.1 added, and (c) estimated K distributions, shown along profile B6–B3; (d–e) simulated concentration after 9.3 and 12 d observed along profiles A1–B3 and B1–A1–B4, respectively.

spreading of the tracer, (2) the importance of possible high ϕ lenses (some visible on logs between 838 and 834 m in wells B3 and A1, including lenses with $\phi > 0.30$) that may be difficult to simulate adequately as they are at the upper end of the ϕ distribution and may have greater occurrence than shown in the available data and global statistics of the site. Uncertainties in these areas may be resolved through future density-driven field, laboratory, and numerical studies and through coupling the hydrogeophysical ϕ estimation and the tracer data inversion together to ensure that a significant near-well structure is adequately treated.

7. Summary and Conclusions

[48] The objective of this study was to investigate several fundamental issues related to a tracer test performed in 2001 at the BHRS including density effects. Our interest was to improve the understanding of (1) hydrologic parameters influencing such transport processes, (2) structural and hydrological parameter information that can be gained from tracer concentration and head data and from other sources of geophysical information and how they may be consistent or complementary, and (3) how different parameter estimation methods may allow various results from the information contained in the data.

[49] With respect to the findings, this study has shown that, at the BHRS at least, ϕ has an important role in the transport and sinking of solute. In fact, through the estimation methods used in this study it is clear that choices made about how the ϕ distribution is defined or included in the estimation approach can produce strong differences in the estimated K distribution. In addition, if ϕ is defined as a prior, some values preclude finding a solution with acceptable fit to the measurements. This finding corroborates results from a recent synthetic study [*Hu et al.*, 2009], showing that variation in ϕ can significantly influence solute plume development.

[50] Furthermore, although hydrological experiments, such as the one investigated here, are known to be relatively ill-posed for the estimation of hydrological property distributions, this is rarely examined in the literature. This study has shown how choices about constraints (assumptions, parameterization, or regularization), which are commonly used to limit effects of ill-posedness in the estimation procedure, can significantly influence the estimated hydrological parameters. Also, for such ill-posed problems where non-uniqueness can be expected, the comparison of multiple estimation approaches and parameterizations can be useful for evaluating which inferred hydrological structures may be realistic, and the extent to which hydrological structures with different scale and/or shape may allow comparable fits of measured BTCs.

[51] With respect to the use of geophysical information to improve the simulation of hydrological processes, here similar hydrological structures seen in the results from different parameter estimation approaches are in good agreement with structures in cross-hole GPR velocity tomograms and ϕ logs. Some similarity between GPR time-lapse images and simulated concentration at various times has also been observed. Also, this study has shown again that the use of geophysical or other additional information is a key to potentially reducing the ill-posedness present in the hydrological parameter estimation process. In this context, the structure imaged from geophysical data can clearly help to constrain the solution domain. With regard to constraining the tracer data inversion with prior estimated ϕ distribution, the strong influence of

the hydrological structure and processes around the injection well for early spreading of the plume require further field, laboratory, and numerical investigations to improve our understanding of such processes and further investigation of how to ideally couple the ϕ estimation process with the tracer data inversion.

[52] Finally, the analysis of this tracer test data may bring some new insights about hydrological unit and property distribution. The incomplete fitting of BTCs, which in fact, is quite common to some degree in field studies, may be because of difficulties in very accurately modeling the transport process, the boundary conditions, and/or the details of heterogeneity of hydrological parameter distributions. Nevertheless, on the basis of the inversion results here and on high-porosity lenses observed in ϕ logs, the presence of a relatively higher K and ϕ zone below 838 m compared to above 838 m seems plausible. At this point, the results from multilevel slug tests performed very recently [*Cardiff et al.*, 2011], show consistently higher K between 836 and 838 m than between 838 m and 839.5 m, although the obtained K values are higher than observed in this study. The same study, *Cardiff et al.* [2011] shows that the commonly used assumption of a strong correlation between ϕ and K in unconsolidated coarse sedimentary aquifers may not be fully applicable at the BHRS. This, as well as comparison and interpretation of the various hydrogeological structures recognized at the BHRS will be the subject of future research. In a more general context, insights into the use of the methods examined here can hopefully help future tracer test experiments with choices in acquisition and processing, strategies for multiple geophysical and hydrological measurement acquisitions, and choices made during modeling and inversion of the experimental data.

[53] **Acknowledgments.** This research was supported by funding to B. Dafflon from the Swiss National Science Foundation and Boise State University. W. Barrash and M. Cardiff were supported by the EPA under grants X-96004601-0 and X-96004601-1, and by the U. S. RDECOM ARL Army Research Office under grant W911NF-09-1-0534. The authors would like to thank the various researchers responsible for collecting and archiving the BHRS data sets used in this analysis. In addition, the authors wish to thank the helpful comments of Walter Illman, Junfeng Zhu, and one anonymous reviewer, whose input helped improve this publication. Product and manufacturer identification are provided for information purposes only.

References

Aster, R. C., B. Borschers, and C. H. Thurber (2005), *Parameter Estimation and Inverse Problems*, 301 pp., Academic, N. Y.

Barrash, W., and T. Clemo (2002), Hierarchical geostatistics and multifacies systems: Boise Hydrogeophysical Research Site, Boise, Idaho, *Water Resour. Res.*, *38*(10), 1196, doi:10.1029/2002WR001436.

Barrash, W., and E. C. Reboulet (2004), Significance of porosity for stratigraphy and textural composition in subsurface, coarse fluvial deposits: Boise Hydrogeophysical Research Site, *Geological Society of America Bulletin*, *116*(9–10), 1059–1073.

Barrash, W., T. Clemo, D. Hyndman, E. C. Reboulet, and E. Hausrath (2002), Tracer/time-lapse radar imaging test: Design, operation, and preliminary results, *Tech. Rep.*, 120 pp., CGISS, Boise State Univ., ID, available at http://cgiss.boisestate.edu/TTLIT/TTLT_report.pdf.

Barrash, W., T. Clemo, J. J. Fox, and T. C. Johnson (2006), Field, laboratory, and modeling investigation of the skin effect at wells with slotted casing, Boise Hydrogeophysical Research Site, *J. Hydrol.*, *326*(1–4), 181–198.

Beinhorn, M., P. Dietrich, and O. Kolditz (2005), 3-D numerical evaluation of density effects on tracer tests, *J. Contam. Hydrol.*, *81*(1–4), 89–105.

Brauchler, R., R. Liedl, and P. Dietrich (2003), A travel time based hydraulic tomographic approach, *Water Resour. Res.*, *39*(12), 1370, doi:10.1029/2003WR002262.

Brauchler, R., R. Hu, P. Dietrich, and M. Sauter (2011), A field assessment of high-resolution aquifer characterization based on hydraulic travel time and hydraulic attenuation tomography, *Water Resour. Res.*, *47*, W03503, doi:10.1029/2010WR009635.

Cardiff, M., W. Barrash, P. K. Kitanidis, B. Malama, A. Revil, S. Straface, and E. Rizzo (2009), A potential-based inversion of unconfined steady-state hydraulic tomography, *Ground Water*, *47*(2), 259–270.

Cardiff, M., W. Barrash, M. Thoma, and B. Malama (2011), Information content of slug tests for estimating hydraulic properties in realistic, high-conductivity aquifer scenarios, *J. Hydrol.*, *403*(1–2), 66–82.

Carrera, J., A. Alcolea, A. Medina, J. Hidalgo, and L. J. Slooten (2005), Inverse problem in hydrogeology, *Hydrogeol. J.*, *13*(1), 206–222.

Certes, C., and G. Demarsily (1991), Application of the pilot point method to the identification of aquifer transmissivities, *Adv. Water Resour.*, *14*(5), 284–300.

Christakos, G. (2000), *Modern Spatiotemporal Geostatistics*, 288 pp., Oxford Univ. Press, N. Y.

Cvetkovic, V., H. Cheng, J. Byegard, A. Winberg, E. L. Tullborg, and H. Widestrand (2010), Transport and retention from single to multiple fractures in crystalline rock at Äspö (Sweden): 1. Evaluation of tracer test results and sensitivity analysis, *Water Resour. Res.*, *46*, W05505, doi:10.1029/2009WR008013.

Dafflon, B., J. Irving, and K. Holliger (2009), Simulated-annealing-based conditional simulation for the local-scale characterization of heterogeneous aquifers, *J. Appl. Geophys.*, *68*(1), 60–70.

Dafflon, B., J. Irving and K. Holliger (2010), Calibration of high-resolution geophysical data with tracer test measurements to improve hydrological predictions, *Adv. Water Resour.*, *33*(1), 55–68.

Dafflon, B., J. Irving, and W. Barrash (2011), Inversion of multiple intersecting high-resolution crosshole GPR profiles for hydrological characterization at the Boise Hydrogeophysical Research Site, *J. Appl. Geophys.*, *73*, 305–314.

de Marsily, G., J. P. Delhomme, F. Delay, and A. Buoro (1999), 40 years of inverse problems in hydrogeology, *Comptes rendus de l'academie des sciences*, *329*(2), 73–87.

de Marsily, G., F. Delay, J. Goncalves, P. Renard, V. Teles, and S. Violette (2005), Dealing with spatial heterogeneity, *Hydrogeol. J.*, *13*(1), 161–183.

Deutsch, C. V., and A. G. Journel (1998), *GSLIB: Geostatistical Software Library and User's Guide*, 384 pp., Oxford Univ. Press, N. Y.

Diersch, H. J. G., and O. Kolditz (2002), Variable-density flow and transport in porous media: Approaches and challenges, *Adv. Water Resour.*, *25*(8–12), 899–944.

Doherty, J. (2003), Ground water model calibration using pilot points and regularization, *Ground Water*, *41*(2), 170–177.

Guo, W., and C. D. Langevin (2002), User's guide to SEAWAT: A computer program for simulation of three-dimensional variable-density ground-water flow techniques of Water Resources Investigations, Book 6, Chapt. A7, p. 77, U.S. Geol. Surv., Reston, Va.

Harbaugh, A. W., E. R. Banta, M. C. Hill, and M. G. McDonald (2000), MODFLOW-2000, The U.S. Geological Surv. modular ground-water model—User guide to modularization concepts and the ground-water flow process, Vol. Open File Rep. 00-92, p. 130., U.S. Geol. Surv., Reston, Va.

Hausrath, E., W. Barrash, and E. C. Reboulet (2002), Water sampling and analysis for the Tracer/Time-Lapse Radar Imaging Test at the Boise Hydrogeophysical Research Site: *Tech. Rep.,* 86 pp., CGISS, Boise State University, ID, available at http://cgiss.boisestate.edu/TTLIT/Chem_report.pdf.

Hu, B. X., M. M. Meerschaert, W. Barrash, D. W. Hyndman, C. M. He, X. Y. Li, and L. J. Guo (2009), Examining the influence of heterogeneous porosity fields on conservative solute transport, *J. Contam. Hydrol.*, *108*(3–4), 77–88.

Hubbard, S. S., J. S. Chen, J. Peterson, E. L. Majer, K. H. Williams, D. J. Swift, B. Mailloux, and Y. Rubin (2001), Hydrogeological characterization of the South Oyster Bacterial Transport Site using geophysical data, *Water Resour. Res.*, *37*(10), 2431–2456.

Hyndman, D. W., and S. M. Gorelick (1996), Estimating lithologic and transport properties in three dimensions using seismic and tracer data: The Kesterson aquifer, *Water Resour. Res.*, *32*(9), 2659–2670.

Hyndman, D. W., J. M. Harris, and S. M. Gorelick (2000), Inferring the relation between seismic slowness and hydraulic conductivity in heterogeneous aquifers, *Water Resour. Res.*, *36*(8), 2121–2132.

Illman, W. A., X. Liu, S. Takeuchi, T.-C. J. Yeh, K. Ando, and H. Saegusa (2009), Hydraulic tomography in fractured granite: Mizunami Underground Research site, Japan, *Water Resour. Res.*, *45*, W01406, doi:10.1029/2007WR006715.

Johnson, B. (2011), Evapotranspiration in the riparian zone of the lower Boise River with implications for groundwater flow, M.S. thesis, Boise State Univ., ID.

Johnson, T. C., P. S. Routh, W. Barrash, and M. D. Knoll (2007), A field comparison of Fresnel zone and ray-based GPR attenuation-difference tomography for time-lapse imaging of electrically anomalous tracer or contaminant plumes, *Geophysics*, *72*(2), G21–G29.

Johnson, T. C., R. J. Versteeg, H. Huang, and P. S. Routh (2009), Data-domain correlation approach for joint hydrogeologic inversion of time-lapse hydrogeologic and geophysical data, *Geophysics*, *74*(6), F127–F140.

Langevin, C. D., W. B. Shoemaker, and W. Guo (2003), MODFLOW-2000, the U.S. Geol. Surv. modular ground-water model–documentation of the SEAWAT-2000 version with the variable-density flow process (VDF) and the integrated MT3DMS transport process (IMT), p. 43., *U.S. Geol. Surv. Open-File Rep. 03-426*, Reston, Va.

Langevin, C. D., D. T. Thorne Jr., A. M. Dausman, M. C. Sukop, and W. Guo (2007), SEAWAT version 4: A computer program for simulation of multi-species solute and heat transport, Vol. 6, Chapt. A22, p. 39, U.S. Geol. Surv. Techniques and Methods, Reston, Va.

Leblanc, D. R., S. P. Garabedian, K. M. Hess, L. W. Gelhar, R. D. Quadri, K. G. Stollenwerk, and W. W. Wood (1991), Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts. 1. Experimental-design and observed tracer movement, *Water Resour. Res.*, *27*(5), 895–910.

Linde, N., S. Finsterle, and S. Hubbard (2006), Inversion of tracer test data using tomographic constraints, *Water Resour. Res.*, *42*(4), W04410, doi:10.1029/2004WR003806.

Mackay, D. M., D. L. Freyberg, and P. V. Roberts (1986), A natural gradient experiment on solute transport in a sand aquifer. 1. Approach and overview of plume movement, *Water Resour. Res.*, *22*(13), 2017–2029.

Malama, B., and B. Johnson (2010), Analytical modeling of saturated zone head response to evapotranspiration and river-stage fluctuations, *J. Hydrol.*, *382*(1–4), 1–9.

Menke, W. (1984), *Geophysical Data Analysis: Discrete Inverse Theory*, 289 pp., Academic, N. Y.

Müller, K., J. Vanderborght, A. Englert, A. Kemna, J. A. Huisman, J. Rings, and H. Vereecken (2010), Imaging and characterization of solute transport during two tracer tests in a shallow aquifer using electrical resistivity tomography and multilevel groundwater samplers, *Water Resour. Res.*, *46*, W03502, doi:10.1029/2008WR007595.

Mwenifumbo, C. J., W. Barrash, and M. D. Knoll (2009), Capacitive conductivity logging and electrical stratigraphy in a high-resistivity aquifer, Boise Hydrogeophysical Research Site, *Geophysics*, *74*(3), E125–E133.

Nelson, G. K. (2007), Deterministic modeling of bromide tracer transport during the tracer/time-lapse radar imaging test at the Boise Hydrogeophysical Research site, M.S. thesis, Boise State Univ., ID.

Pollock, D., and O. A. Cirpka (2010), Fully coupled hydrogeophysical inversion of synthetic salt tracer experiments, *Water Resour. Res.*, *46*(7), W07501, doi:10.1029/2009WR008575.

Ptak, T., and G. Schmid (1996), Dual-tracer transport experiments in a physically and chemically heterogeneous porous aquifer: Effective transport parameters and spatial variability, *J. Hydrol.*, *183*(1–2), 117–138.

Ramarao, B. S., A. M. Lavenue, G. Demarsily, and M. G. Marietta (1995), Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields. 1. Theory and computational experiments, *Water Resour. Res.*, *31*(3), 475–493.

Rubin, Y. (1995), Flow and transport in bimodal heterogeneous formations, *Water Resour. Res.*, *31*(10), 2461–2468.

Rubin, Y., X. Y. Chen, H. Murakami, and M. Hahn (2010), A Bayesian approach for inverse modeling, data assimilation, and conditional simulation of spatial random fields, *Water Resour. Res.*, *46*, W10523, doi:10.1029/2009WR008799.

Sambridge, M., and K. Mosegaard (2002), Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.*, *40*(3), 1009, doi:10.1029/2000RG000089.

Scheibe, T. D., and Y. J. Chien (2003), An evaluation of conditioning data for solute transport prediction, *Ground Water*, *41*(2), 128–141.

Simmons, C. T. (2005), Variable density groundwater flow: From current challenges to future possibilities, *Hydrogeol. J.*, *13*(1), 116–119.

Simmons, C. T., T. R. Fenstemaker, and J. M. Sharp (2001), Variable-density groundwater flow and solute transport in heterogeneous porous media: Approaches, resolutions and future challenges, *J. Contam. Hydrol.*, *52*(1–4), 245–275.

Singha, K., and S. M. Gorelick (2006), Hydrogeophysical tracking of three-dimensional tracer migration: The concept and application of apparent petrophysical relations, *Water Resour. Res.*, *42*(6), W06422, doi:10.1029/2005WR004568.

Slater, L. D., A. Binley, and D. Brown (1997), Electrical imaging of fractures using ground-water salinity change, *Ground Water*, *35*(3), 436–442.

Stein, M. (1987), Large sample properties of simulations using Latin hypercube sampling, *Technometrics*, *29*(2), 143–151.

Sudicky, E. A. (1986), A natural gradient experiment on solute transport in a sand aquifer–spatial variability of hydraulic conductivity and its role in the dispersion process, *Water Resour. Res.*, *22*(13), 2069–2082.

Sudicky, E. A., W. A. Illman, I. K. Goltz, J. J. Adams, and R. G. McLaren (2010), Heterogeneity in hydraulic conductivity and its role on the macroscale transport of a solute plume: From measurements to a practical application of stochastic flow and transport theory, *Water Resour. Res.*, *46*, W01508, doi:10.1029/2008WR007558.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, 352 pp., Society for Industrial and Applied Mathematics, Philadelphia, Pa.

Vereecken, H., U. Doring, H. Hardelauf, U. Jaekel, U. Hashagen, O. Neuendorf, H. Schwarze, and R. Seidemann (2000), Analysis of solute transport in a heterogeneous aquifer: The Krauthausen field experiment, *J. Contam. Hydrol.*, *45*(3–4), 329–358.

Yeh, T. C. J., and S. Y. Liu (2000), Hydraulic tomography: Development of a new aquifer test method, *Water Resour. Res.*, *36*(8), 2095–2105.

Yeh, W. W. G. (1986), Review of parameter-identification procedures in groundwater hydrology: The inverse problem, *Water Resour. Res.*, *22*(2), 95–108.

Zheng, C. (2005), MT3DMS v5 supplemental user's guide: Dept. of Geological Sciences, Univ. Alabama, *Tech. Rep. to the U.S. Army Engineer Research and Development Center*, p. 24.

Zheng, C., and P. P. Wang (1999), MT3DMS: A modular three-dimensional multispecies transport model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems, documentation and user's guide, *Vol. Contract Rep. SERDP – 99 – 1*, U.S. Army Corps of Engineers, Engineer Research and Development Center, p. 220.

Zhu, J. F., and T. C. J. Yeh (2005), Characterization of aquifer heterogeneity using transient hydraulic tomography, *Water Resour. Res.*, *41*(7), W07028, doi:10.1029/2004WR003790.

W. Barrash and M. Cardiff, Center for Geophysical Investigation of the Shallow Subsurface, Boise State University, 1910 University Dr., Boise, Idaho 83725, USA. (baptiste.dafflon@gmail.com)

B. Dafflon, Now at Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA.

T. C. Johnson, Pacific Northwest National Laboratory, 902 Battelle Blvd., Richland, WA 99352, USA.